DOCUMENT RESUME

ED 030 486 PS 001 938

By-Meyer, William J.; And Others

Project Head Start Research and Evaluation Center, Syracuse University, Research Institute. Final Report.

Syracuse Univ., N.Y. Research Inst.

Spons Agency-Office of Economic Opportunity, Washington, D.C.

Report No-OEO-4120

Pub Date 31 Aug 68

Note-114p.

EDRS Price MF-$0.50 HC-$5.80

Descriptors-*Annual Reports, Behavioral Science Research, *Behavior Rating Scales, Cognitive Development, Concept Formation, *Discrimination Learning, Educational Research, Factor Analysis, Factor Structure, *Intelligence Tests, *Peer Relationship, Preschool Children, Program Evaluation

Identifiers-Adaptive Behavior Rating Scale, *Head Start, Stanford Binet

The following research projects are described in this annual report: (1) "Concept Learning in Discrimination Tasks," which indicates that kindergarten children are able to discriminate the letters "b," "d," "p," and "q"; (2) "Discrimination of Letter-like Forms," indicating that nursery school children discriminate as well as second graders in brief but appropriate experiences; (3) "Presence of an Attached Peer and Security in a Novel Environment," which supports the hypothesis that proximity of a peer friend has a distress inhibiting or security inducing effect on children which is greater than that afforded by proximity to a strange peer; (4) "Kindergarten 'Learning to Learn' Program Evaluation," a study of alternative learning situations for effects on young children's development, (5) "Changes in Stanford-Binet IQ Performance vs. Competence," which contends that instead of concentrating on IQ gains in evaluating preschool programs, investigators should study characteristics of children as they interact with characteristics of the program, (6) "The Adaptive Behavior Rating Scale," indicating that crucial adaptive behaviors, as defined by preschool teachers, involve social competency and compliance; and (7) "Dissemination Activities." (MS)

ED030486

F-OEO
OEO - 4120

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

# Project Head Start Research and Evaluation Center
Contract # OEO 4120
Final Report
August 31, 1968

William J. Meyer
Center Director

Cooperating Agency
Syracuse University
Research Institute
201 Marshall Street
Syracuse, New York 13210

Institutional Grants Officer
L. William Patchen
Assistant Director
Research Institute
Syracuse University

PS001938

# TABLE OF CONTENTS

I. Research Projects

1. Concept Learning in Discrimination Tasks — Page 1
   Edward Caldwell and Vernon C. Hall

2. Discrimination of Letter-like Forms — Page 9
   Vernon C. Hall and Edward Caldwell

3. Presence of an Attached Peer and Security in a Novel Environment — Page 19
   J. Conrad Schwarz

4. Kindergarten "Learning to Learn" Program Evaluation — Page 63
   Margaret Lay

5. Changes in Stanford-Binet IQ. Performance vs. Competence — Page 80
   William J. Meyer and Byron Egeland

6. The Adaptive Behavior Rating Scale — Page 95
   William J. Meyer and Lois Hayweiser

7. Dissemination Activities — Page 106

II Appendix

1. Teacher-Guided Sequentially Arranged Program — Page 109

2. Expressive Activities Program — Page 110

III. Acknowledgements — Page 112

# Concept Learning in Discrimination Tasks

Edward C. Caldwell and Vernon C. Hall

Syracuse University

The problem of false positives in psychological research has been well demonstrated by Rosenthal (1963) in his discussions of experimenter bias. Campbell (1957) and Campbell and Stanley (1963) have discussed this problem in terms of internal and external validity of experimental designs. Eight different classes of variables related to internal validity, e.g., history, maturation, testing, etc., and four classes of variables related to external validity, e.g., interaction of testing and treatment, etc. are presented. These classes of variables serve as plausible rival hypotheses, i.e., rival to the hypothesis one is probing. If these variables are not recognized as sources of rival hypotheses and consequently left uncontrolled, one runs the risk of generating false positives.

While Campbell (1957) and Campbell and Stanley (1963) have presented a fairly complete analysis of the sources of false positives (Rosenthal's experimenter bias may be considered as a case of their instrumentation.), it is the thesis of the present research that an additional and frequently overlooked class of variables can be found in some of the most respected psychological literature which serves as a source of plausible rival hypotheses. This class of variables might be labeled "incomplete task analysis" and is relevant to human learning tasks, especially those concerned with learning in early childhood.

Edward Caldwell's address is: Psychology Department, Syracuse University. This research was supported by grant (OEO 4120) from the U.S. Office of Economic Opportunity.

The field of computer simulation has focused attention on the need for complete task analysis. Feigenbaum and Feldman (1963) have discussed this variable of completeness. In the course of writing a program, i.e., constructing a model of behavior, "the researcher realizes the inadequacy of his original idea(s). He discovers that more information is needed. To get this information he may have to reanalyze old data or run new experiments. Thus the researcher encounters one of the advantages of computer models--the requirement for completeness and precision." "The only real constraint on the model builder is that his statements be unambiguous and complete." It is a very simple task to test a computer model for precision and completeness. Since the model is the program, one need merely run it to check both criteria. Unfortunately, verbal models are not so readily evaluated and thus may often represent incomplete analyses of behavior. Hence, false positives may well develop.

Several studies (Davidson, 1935; Hildreth, 1932; Ilg and Ames, 1950; Smith, 1928; Wilson and Flemming, 1938) have indicated that discrimination of letters that have rotations or reversals is difficult for children typically until ages seven to eight. Vernon (1960) has suggested, "This development seemed due to maturation rather than learning to read, for Hildreth found that among children he studied those who had begun to read showed little superiority to those who had not."

However, studies by Hendrickson and Muehl (1962) and Jeffrey (1958, 1966) have demonstrated that children, when given appropriate experiences, can in fact discriminate not only reversals but also oblique lines. These studies, which use motor responses and arrows in pretraining experiences, suggest that the crucial variable is attentional.

While attention to stimuli is obviously a necessary condition for correct performance on discrimination tasks, it is argued here that these tasks have typically been inadequately analyzed. If one were to program a computer to do any of these discrimination tasks, an appropriate concept of _same_ and _different_ would be essential. No amount of attention to the stimuli will produce a "correct" response in a discrimination task unless the subject holds the same concept of _same_ and _different_ as that demanded by the experimental situation. In the studies cited above this means that orientation differences mean _different_. The studies by Hendrickson and Muehl (1962) and Jeffrey (1958, 1966) involved feedback to the subject which provided the information with which he was able to learn the "correct" concept of _same_ and _different_.

The present research, a variation of Davidson's (1935) study, is the second of a series designed to explore the influence of concept learning on discrimination tasks. (First study see Hall, Caldwell and Simpson, 1968) Davidson described her criterion task, a Letter Perception Test (LPT) thus: "A certain letter was printed in large, black type at the left side of a sheet of paper. After it a 'box' was drawn containing four rows of ten letters each. Among these letters were five identical to the given letter, five of the letter most likely to be confused with it (the mirrored opposite in the case of _d_, _q_ and _b_) several each of possible confusing letters including several tall letters in case 'tallness' was an important factor in letter recognition, while the remainder were randomly selected letters. The children were instructed to look at the given letter, then find everyone exactly like it in the box and mark it when they find it."

One of her well known conclusions was that "it was not until a chronological age of seven and one-half years was reached that 50 per cent of the pupils... were able to select _d_, _q_ and _b_ without error." She further argued, "one cannot but condlude that growth in ability to distinguish between the letters of the alphabet is closely allied with increasing mental age..." "These data indicate that some letters of the alphabet are more difficult to discriminate than others. The letters that are the most difficult to distinguish are the pairs of letters that are the reversal of each other. Letters that are upside down inversions of others are also difficult to differentiate but not nearly so difficult as reversals. Increase in ability to differentiate comes with increasing mental maturity and experience."

While the authors do not quarrel with the Davidson data, (Strikingly similar results were obtained by Hall, Caldwell and Simpson, 1968), they do challenge her interpretations on the basis of an inadequate analysis of the criterion task. When a child marked a _d_ and the standard was a _b_, he might have done so because the letters were indistinguishible to him or because by his definition of _same_ and _different_ this reversal was immaterial. Just as three-dimensional objects go by the same name regardless of their orientation, so would two-dimensional graphemes for the subject unless he had learned a new concept of _same_ and _different_ for this type of stimuli. The argument here is that the child enters the experiment with a concept of _same_ and _different_ for which orientation is irrelevant--a chair is a chair regardless of its orientation--and is expected to function as though orientation were relevant--a _d_ is not a _b_, without any new concept learning. Reversal and rotation confusions would be expected to decline with age since this "new concept learning" would likely occur through experiences with letters and other symbols in formal schooling.

The present experiment was designed to test the following hypotheses:

1. When kindergarten children are given warm-up and feedback on a a task designed to produce a concept of _same_ and _different_ in which orientation is

relevant, they will make fewer confusions on a Letter Perception Test than an appropriate control group.

2. When kindergarten children are given warm-up and feedback on a task designed to produce a concept of *same* and *different* in which orientation is irrelevant, they will make more confusion errors on a Letter Perception Test than an appropriate control group.

## Subjects

Subjects were 72 kindergarten children from Jamesville Elementary School, Jamesville, New York, randomly assigned to one of three treatment groups.

## Method

For all subjects a criterion task similar to Davidson's LPT was used. The standards were *b*, *d*, *p* and *q*. The 40 letters from which these were to be selected included five of each of the standards, five of the letter *h* and 15 randomly assigned letters.

The warm-up task consisted of an overlay procedure and a paper and pencil task, the latter similar to Davidson's LPT. For the overlay procedure the standards in Figure 1 were reproduced on 3M Projection Transparencies 127 and placed in 2" x 2" Super Easymount slide holders for ease of handling. Each set of standards and transformations from Figure 1 was duplicated, cut into 1-11/16" squares, mounted on 7" x 14" pieces of red poster board in a layout similar to Figure 1 and coated with plastic.

*S* was seated at a table with the board before him. *E* instructed *S* to find the ones on the right that were exactly the same as the standard. He further indicated that they were exactly alike if the plastic overlay lined up with the figure when the overlay was placed on top, and that they were different if the lines of the figure showed when the overlay was on top.

Subjects in Group 1 were cautioned not to turn the overlay around or over in any way when trying to line them up, because if they had to turn it in any way, then this meant they were different. Group 2, however, was encouraged to turn the overlay in any way, and as long as they could in some way be lined up, they were the same. Group 3 was given no information regarding turning of the overlay since orientation was irrelevant for the figures used for them. However, since a 25 per cent size differential was used for Group 3, which was difficult to discriminate without careful observation, attention was essential for obtaining correct responses.

The paper and pencil task was presented on two sheets of paper with five sets of standards and transformations on the first sheet and one set at the top of the LPT sheet. Figure 2 shows the standards and transformations used. S was instructed to circle those in the big box that were exactly like the standard. Feedback was given on this task after completion of each standard and transformations set. While transparencies were available for checking S's answers, they were not used by S during this part of the warm up. Rather, S was encouraged to pretend that he was picking up the standard and comparing it like he had done during the overlay task. E then used the transparencies to check S's answers. S was told to do the LPT just like he had done the others. No feedback was given during the LPT.

## Results

The results were analyzed in terms of the per cent of correct responses for each subject on all letters of the LPT. An analysis of variance showed significant treatment effects, F (2, 69) = 35.2; p .01. A comparison of groups using a Scheffe showed that Group 1 performed significantly better than Group 3, F (1, 69) = 27.2; p .01; and Group 3 performed significantly better than Group 2;

PS 001938

$F\ (1, 69) = 11.7;\ p$ .01.

A comparison of the present results with Davidson's in terms of per cent of subjects for the same age range who made no confusion errors is shown in Table 1. Note that Group 3 performed much like Davidson's subjects while Group 1 far out performed the other two groups.

Table 2 shows the per cent of irrelevant errors of the total number of letters circled for each group. These included the letters *h*, *a*, *g*, *c*, etc.

### Discussion

The authors have argued that inappropriate task analysis is a potential source of false positives in research with children. We have suggested that the failure to consider the adequacy of the child's concept of same and different has led several experimenters along with a large number of educators and developmental psychologists into false generalizations concerning the discrimination ability of young children. At the very least, the results clearly indicate that kindergarten children are able to discriminate the letters *b*, *d*, *p* and *q*.

An alternative explanation of the results might revolve around the possibility that the group differences reflect differential degrees of attention required during the warm-up tasks. Thus Group 1 had to attend to both shape and orientation whereas Group 2 had to attend to shape alone. Further, one might argue that Group 3 had to attend closer than Group 2 since for the former, size was relevant as well as shape. (Perhaps size and shape require less attention than orientation and shape.) One way to assess the attention factor is to consider the per cent of irrelevant errors. If attention is used to explain the superiority of group 1 over 3 and 3 over 2, then one would expect equivalent differences in the per cent of irrelevant errors. As can be seen in Table 2, this was not the case.

Thus while attention is obviously an important factor in discrimination learning, it alone does not account for the differences obtained in the present experiment. Rather, if one considers the element of concept learning as a necessary prerequisite for discrimination, the data are readily understood. Subjects in Group 3, when given no information regarding the relevance of orientation, apparently utilized the concept of same and different they brought to the experimental situation, and, consequently, performed much like the S's in Davidson's study. Subjects in Group 2 were outperformed by Group 3 because the former were operating on a concept of same and different, learned during the warm-up, which was inappropriate for the criterion task. Group 1, however, was given correct information regarding the relevance of orientation in the two-dimensional world of graphemes, and thus performed with relatively few errors on the LPT. Hence with more adequate task analysis and the implementation of required steps in the learning procedure, children are able to perform much more adequately.

## Bibliography

Campbell, D.T. Factors relevant to the validity of experiments in social settings. Psychol. Bull., 1957, 54, 297-312.

Campbell, D.T. and Stanley, J.C. Experimental and quasi-experimental designs for research in teaching. In Gage, N.L. (Ed.) Handbook of Research on Teaching. Skokie, Ill., Rand McNally & Co. 1963.

Davidson, H.P. A study of the confusing letters B,D,P and Q. J. Genet. Psychol. 1935, 47, 458-468.

Feigenbaum, E.A. and Feldman, Julian Computers and Thought, New York: McGraw-Hill Co. 1963.

Hall, Vernon C., Caldwell, E.C. and Simpson, G. Variables affecting the performance of young children on a letter discrimination task. Paper read at AERA, Chicago, 1968.

Hendrickson, L.N. and Muehl, S. The effect of attention and motor responses pretraining on learning to discriminate B and D in kindergarten children. J. Educ. Psychol., 1962, 53, 236-241.

Jeffrey, W.E. Variables in early discrimination learning: I. Motor responses in training a left-right discrimination. Child Developm. 1958, 29, 269-275.

Jeffrey, W.E. Discrimination of oblique lines by children. J. comp. physiol. Psychol. 1966, 62, 154-156.

Ilg, F.L. and Ames, L.B. Developmental trends in reading behavior. J. genet. Psychol. 1950, 76, 291-312.

Rosenthal, Robert On the social psychology of the psychological experiment: The experimenter's hypothesis as unintended determinant of experimental results. Amer. Sci. 1963, 51, 268-283.

Smith, N.B. Matching ability and first grade reading. J. educ. Psychol., 1928, 19, 560-571.

Vernon, M.D. Backwardness in Reading. Cambridge: Cambridge University Press, 1957.

Wilson, F.T. and Flemming, C.W. Reversals in reading and writing. J. genet. Psychol. 53, 1938, 3-31.

# DISCRIMINATION OF LETTER-LIKE FORMS

Vernon C. Hall and Edward Caldwell

## Analysis

The developmental curves of the original Gibson, Gibson, Pick and Osser (1962) publication have appeared in several prominent places, (Epstein, 1967; Gibson, 1965; Gibson, 1963a, 1963b).

In her article in *Science*, Gibson discusses the significance of these curves for teaching of reading. She argues that "some transformations are harder to discriminate than others and improvement occurs at different rates for different transformations." Gibson, 1965, p. 1067). The question of how this improvement occurs is of significance for reading readiness and instruction. Gibson suggests, "that it is perceptual learning and need not be verbalized is probable (though teachers do often call attention to contrasts between letter shapes)." (Gibson, 1965, p. 1068). If, as the present author has argued, concept learning is an important element in the slope of these developmental curves, then verbalization may be extremely valuable.

Gibson *et al*. (1962) presented subjects with 12 letter-like forms, each of which had twelve transformations. "The transformations chosen were as follows: three degrees of transformations of line to curve or curve to line (1, 2, or 3); five transformations of rotation or reversal, i.e., 45° rotation, 90° rotation, 180° rotation, right-left reversal, and up-down reversal; two perspective transformations, a 45° slant left and a 45° backward tilt; and two topological transformations, a break and a close." "...half of the forms (the symmetrical ones) lacked two transformations, since either the right-left reversal or up-down reversal was identical with the standard and the 180° rotation was identical with the other. The cells in these cases were filled with the standard." (Gibson, 1962, pp. 897-898.)

A standard form was placed on the top center row of a five row matrix board. Each of the four rows below the standard held 13 cards (1 ¼" square on which the transformations were mounted). In each of these rows a given standard and its twelve transformations were placed in random order. "After E had put the appropriate standard in the center of the top row, S went through a given row searching for any form which was 'exactly like the standard.' When he found one he removed it and handed it to E and continued until he had scanned the entire row." Gibson (1962, p. 898.)

The subjects were from 4 to 8 years old. Errors were scored according to four categories of 1) perspective, 2) rotation and reversal, 3) line to curve, and 4) break and close. Regarding the decline in confusion errors with age, the authors argue, "Better ability to keep the task in mind and follow it through cannot be wholly ruled out, but preliminary practice, instructions and assistance in scanning when required, rendered this factor minimal. Maturation of retinal processes is still continuing during this period, but acuity as such should have been adequate even for the four year olds. It seems, rather, that children between 4 and 8 learn something about letter-like forms which makes possible better discrimination even between ones they have not seen before."

"Our results suggest that what they learn are the features or dimensions of difference which are critical for differentiating letters."... "It is our hypothesis that it is the distinctive features of grapheme patterns which are responded to in discrimination of letter-like forms. The improvement in such discrimination from four to eight is the result of learning to detect these invariants and of becoming more sensitive to them." Gibson et al. 1962, p. 904. Pick (1965) argues similarly that Gibson et al. have "demonstrated that children between ages of four and eight years improve in their ability to make visual discriminations among letter-like forms." (Pick, 1965, p. 331). (The underlining is the present author's.)

As in the Davidson study discussed in the previous chapter, the behavioral model here is incomplete. An adequate concept of same and different is essential in performing the discrimination task. The differential performance over age groups, it is argued, is not due merely to "learning to detect", "becoming more sensitive to distinctive features", or "improvement in visual discrimination," if to these at all. One may detect and be highly sensitive to distinctive features in the discrimination task, but unless S holds the appropriate concept of same and different, i.e., identical to E's concept, no amount of concern with distinctive features, no amount of visual detection will provide S with a correct response. Detection of distinctive features is no doubt a necessary but not a sufficient condition for success on the discrimination task.

The second experiment was designed to test the following hypothesis:

1. When nursery school children are given warm-up and feedback on a task designed to produce a concept of same and different identical to E's concept, there will be no difference between nursery school children and second graders on the line to curve and rotation and reversal errors. (No differences were found on the other transformations by Gibson et al. )

2. When nursery school children are given warm-up and feedback which give inappropriate information regarding rotation and reversal but give appropriate information regarding line to curve transformations, they will make more rotation and reversal errors than second graders or nursery school children given appropriate information regarding rotations and reversals.

3. When second graders are given warm-up and feedback as in 2 above, they will make more rotation and reversal errors than second graders and nursery school children given appropriate information.

## Subjects

Subjects were 72 nursery school children and 72 second graders. Each grade level was randomly assigned to one of four treatment groups.

## Procedures

The criterion task was identical to that used by Gibson *et al*. The warm-up task was nearly identical to that used in the study discussed in the previous chapter. However, in place of the paper and pencil task, a task similar to Gibson's criterion task was employed. *S* worked from a matrix board and tipped down the appropriate symbols after selection. (Tipping them down rather than handing them to *E* greatly facilitated their replacement for the next subject.) *S* was given feedback consistent with his group designation. Group 1-N was the nursery school children given the same warm-up and information regarding same and different as was given Group 1 in the study discussed in the previous chapter, (orientation relevant) and Group 1-2 was their second grade counterpart. Group 2-N was the nursery school children given the same information as was given Group 2 in the previous study (orientation irrelevant), and Group 2-2 was their second grade counterpart. Group 3-N was the nursery school children given the same information as was given Group 3 in the previous study and Group 3-2 was their nursery school children counterpart. Group 4-N was the nursery school children given warm-up identical to that used by Gibson *et al*. and Group 4-2 was their second grade counterpart.

## Results

The results were analyzed in terms of per cent of stimulus cards chosen in each of five categories: 1) correct, i.e., identical to the standard, 2) line to curve transformations, 3) rotation and reversal transformations, 4) perspective transformations, and 5) break and close transformations. Table 3 gives the mean per cent and standard deviations in each of the eight groups for the five categories. Five 4 x 2 (treatment by grade) analysis of variance were computed, one for each category. Since the nature of the treatments would tend to produce heterogeneity of variances, significance was set at the .01 level. Likewise, for comparison of means, the Scheffe test was used with the level of significance set at .01. A summary of the analyses of variance is presented in Table 4. Table 5 shows the matrix of results from the Scheffe tests with the column heading indicating the group that made more selections.

The predictions from the hypotheses were: 1) Groups 1-N and 1-2 would not differ from each other on line to curve or rotation and reversal errors. Table 5 shows that this hypothesis was confirmed. 2) Group 2-N would make more rotation and reversal errors than Group 1-N and 1-2, and these three groups would not differ on line to curve errors. Table 4 shows that while the predicted differences occurred in the rotation and reversal errors, Group 2-N also made more line to curve errors than Groups 1-N and 1-2. 3) Group 2-2 would make more rotation and reversal errors than Groups 1-N and 1-2, and the three groups would not differ on rotation and reversal errors. Table 4 shows that this prediction was confirmed.

## Discussion

First it should be noted from Table 3 that the nursery school children in the replication study, Group 4-N, performed very much like the 4 year olds in the gibson *et al*. study. Furthermore, Table 5 shows that Group 4-N made more break and close errors than did their second grade counterpart. Since Gibson *et al*.

did not obtain this result, it can be argued that the nursery school children in the present study were certainly not superior to the 4 year olds of the Gibson et al. study at least relative to the respective 8 year olds or second graders. Hence the superiority of Group 1-N cannot likely be attributed to population differences.

There were no differences in the number of correct responses. Hence the treatments did not merely manipulate the quantity of cards selected by S. The fact that there were no differences between Groups 1-N and 1-2 on any of the transformations indicates that nursery school children perform much like second graders when given an eight to ten minute warm-up designed to give them an adequate concept of same and different. Since Groups 2-N and 2-2 made more rotation and reversal errors than all of the other groups, it is obvious that they quickly employed the concept that shape, regardless of orientation, was what E meant by same. Figure 4 clearly indicates this result of concept learning.

The unanticipated result that Group 2-N made more line to curve errors than Groups 1-N and 1-2 requires consideration. During warm-up, since orientation was irrelevant, less attention was demanded of Group 2-N for correct responses. No close inspection of the stimuli was required since those that were different were different on many dimensions. Hence their poorer performance may well have been due to a set of looking only for an object that had roughly the same shape. Attention or in Gibson's terms, "learning to detect," "becoming more sensitive to distinctive features" is an important factor here.

Obviously then both attention and concept learning are important variables in this discrimination task. An attempt was made to separate the contribution of each by use of Group 3-N and 3-2, for which attention was essential (due to the size transformation) but for which relatively little information was available regarding orientation. Since Groups 1-N and 3-N did not differ on any transformation at the .01 level, it can be argued that there was little if any difference in the attentional requirements for the two groups. Further, the difference that was greatest between the two groups was on rotation and reversal errors. This was significant at the .05 level using the Scheffe comparison. Further, the subjects in Group 3-2 performed significantly better than all the other second graders in perspective transformations. This discrimination requires considerable attention. A final suggestion that attention was equal for the two kinds of warm-up conditions comes from the study discussed in the previous chapter where it was seen that irrelevant errors indicated equivalent attention.

It is tempting to view the difference observed in Figure 4 between Groups 1-N and 3-N as the effects of the orientation factor, i.e., concept learning, and the difference between Groups 1-N and 4-N as a combination of orientation and attention factors. Certainly the latter can be accepted.

Regardless of the relative influence of these two variables, an important and obvious conclusion is that nursery school children are able to discriminate as well as second graders when given very brief but appropriate experiences. Figure 5 demonstrates graphically the irrelevance of grade and/or age as a variable since Group 1-N, a group of nursery school children, coincides very closely with the 8 year olds of the Gibson *et al*. study and the present replication group, Group 4-2. Once again it is apparent that when tasks are more adequately analyzed into their component skills and young children are trained in these skills, they perform much more adequately than many of the normative studies indicate they are capable of.

## Table 4

## Analyses of Variance Summaries

| Transformation | Source | df | SS | MS | F |
|---|---|---|---|---|---|
| Correct | Treatments (A) | 3 | 1,551 | 517.0 | 4.58*** |
| | Grade level (B) | 1 | 205 | 205.0 | 1.28 |
| | A x B | 3 | 568 | 189.3 | 1.68 |
| | Error | 136 | 15,341 | 112.8 | |
| Line to Curve | Treatments (A) | 3 | 3,767 | 1,255.7 | 6.3** |
| | Grade level (B) | 1 | 18,067 | 18,067 | 90.8** |
| | A x B | 3 | 2,726 | 908.7 | 4.6** |
| | Error | 136 | 27,057 | 198.9 | |
| Rotation & Reversal | Treatments (A) | 3 | 111,477 | 37,159 | 99.5** |
| | Grade level (B) | 1 | 13,964 | 13,964 | 37.4** |
| | A x B | 3 | 11,811 | 3,937 | 10.5** |
| | Error | 136 | 50,762 | 373.3 | |
| Perspective | Treatments (A) | 3 | 11,436 | 3,812 | 11.6** |
| | Grade level (B) | 1 | 7,951 | 7,951 | 24.2** |
| | A x B | 3 | 5,762 | 1,921 | 5.8** |
| | Error | 136 | 44,738 | 329.0 | |
| Break & Close | Treatments (A) | 3 | 2,153 | 717.7 | 3.4* |
| | Grade level (B) | 1 | 8,289 | 8,289.0 | 39.5** |
| | A x B | 3 | 1,428 | 476.0 | 2.3 |
| | Error | 136 | 28,528 | 209.8 | |

*p .05; **p .01; *** While this was significant, the Scheffe comparison showed no differences. Groups 1-4 and 2-4 were most different. See Table 3.

Table 5

Summary of Scheffe Comparisons of Means

| Groups | 1-4 | 1-8 | 2-4 | 2-8 | 3-4 | 3-8 | 4-4 | 4-8 |
|---|---|---|---|---|---|---|---|---|
| 1-4 | | | L-C<br>R&R | R&R | | | L-C<br>R&R | |
| 1-8 | | | L-C<br>R&R | R&R | R&R | | L-C<br>R&R<br>PER | |
| 2-4 | | | | | | | | |
| 2-8 | | | L-C | | | | L-C<br>PER | |
| 3-4 | | | R&R | R&R | | | | |
| 3-8 | | PER | L-C<br>R&R<br>PER | R&R<br>PER | L-C<br>R&R<br>PER | | L-C<br>R&R<br>B&C<br>PER | PER |
| 4-4 | | | R&R | R&R | | | | |
| 4-8 | | | L-C<br>R&R | R&R | R&R | | L-C<br>R&R<br>B&C | |

Column heading indicates group making more selections. L-C is line to curve; R&R is rotation and reversal; B&C is break and close; PER is perspective.

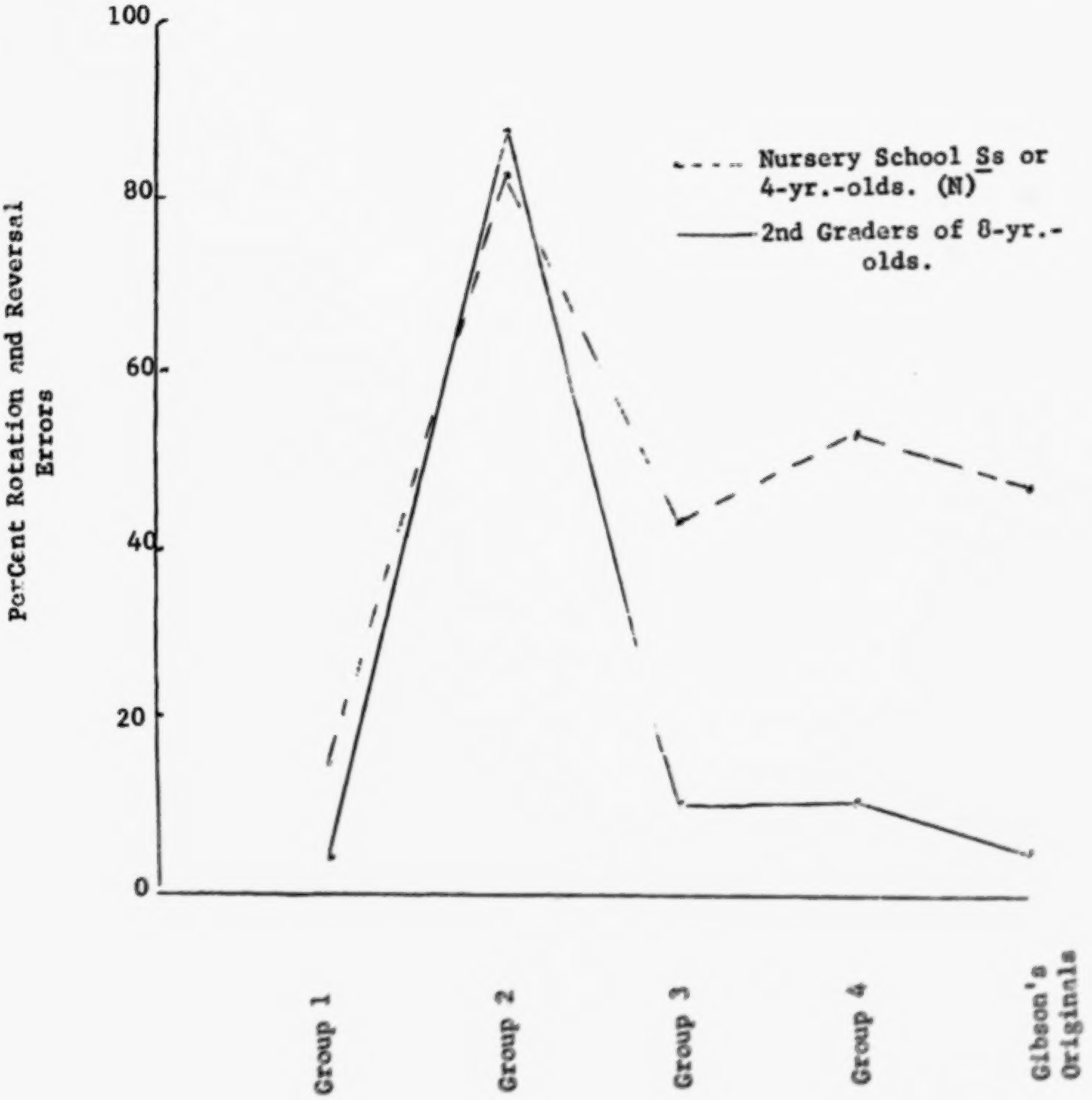

Fig. 4. -- Per cent of rotation and reversal errors by groups.

# Presence of an Attached Peer and Security in a Novel Environment

J. Conrad Schwarz

Syracuse University

In recent years a number of review and theoretical artacles have been written on the development of social attachments (Cairns, 1965; Fraiberg, 1967; Gray, 1958; Harlow and Harlow, 1966; Schaffer and Emerson, 1964; Scott, 1963, 1967). These articles provide evidence that for many mammals, indluding man, exposure to another organism of sufficient duration and intensity and at the appropriate stage in development produces an attachment which is manifested by the maintenance of proximity to the object of attachment and by the occurrence of distress or behavioral disruption upon separation from that organism. A longitudinal study of the development of social attachments in human infants by Schaffer and Emerson (1964) indicated that all infants gave evidence of attachments to specific individuals by 18 months of age. For the great majority, the attachment was formed between six and ten months of age. A social attachment may be thought of as one manifestation of a motivational system which in its typical mode of operation insures the maintenance of proximity to other members of the species. Proximity, in turn, provides the opportunity for learning through modeling or observation and symbolic commumication, and for mutual protection and cooperative endeavor.

In addition to the general consequences cited above, the proximity of attached objects appears to have important effects upon the emotional state of the subject. About the same time that specific attachments are being formed, the human infant begins to show fear of strangers, i.e., about seven to nine months of age (Bronson, 1968a, 1968b). The response to strangers may be just one manifestation of the emergence of a more general fear of the strange, since there is evidence that beginning at about seven months the infant also responds with distress to strange environments (Schaffer and Callender, 1959). Thus it appears that as the infant begins to be disturbed by strange stimulus configurations he also begins to prefer the proximity of specific individuals, a circumstance which presumably affords some relief from fear of the strange. Direct evidence that the mother plays the unique role of inhibiting distress in a novel situation is found in an experimental study by Rheingold (in press). She found that ten month old infants left in an unfamiliar room either alone or with a stranger cried almost immediately and prolongedly, whereas the mother's passive presence in the room resulted in contented play and no crying whatever. Thus the presence of the mother, presumedly an attached object at 10 months of age, inhibited distress in a novel environment. Further evidence for the inhibition of distress by proximity to an attached object in human infants is found in a study by Morgan and Ricciuti (in press). They observed that the approach of a stranger evoked distress in 12 month old infants who were seated in a baby tenda six feet from their mother; however, in another experimental condition, in which the infant was seated on the mother's lap, the distress evoked by the approach of the stranger was significantly attenuated. In more naturalistic settings one year olds have been observed to use the mother as a secure base of operations from which brief exploratory excursions are made into less familiar territory (Ainsworth, 1963; Ainsworth and Wittig, in press).

In a review of maternal-infant relationships in mammals, King (1966) theorized that the distress which would follow exposure of a solitary infant to novel stimulus is inhibited in the mother's presence because her proximity evokes a positive affective state in the infant. Schwarz (1968) has proposed that in the absence of a specific attachment to the mother, her presence will neither evoke a positive affective state nor inhibit distress. (Bronson, 1968a, has reached similar conclusions.) This proposition is consistent with the discrepant behaviors of rhesus monkeys when exposed to a novel environment in the presence of a wire surrogate mother as opposed to a cloth surrogate (Harlow and Zimmerman, 1959). In regard to the cloth mothers, the quality of the stimulation they provided may have been a more important factor in their capacity to inhibit distress than any product of the history of surrogate-infant interaction. On the other hand, the human infant's response to the mother at one year of age is distinctly different than that to a strange female, yet both may afford similar physical stimulation. The difference in the affect which attends proximity to each may lie in differential levels of arousal which each may evoke in the child. King (1966) has suggested that the fear evoking potential of a stimulus configuration is a function of the ratio of novel stimulus elements to total stimulus elements. Approaching a stranger increases the proportion of stimulus elements in the total perceptual field which are novel. Approaching the mother, on the other hand, reduces the total number of effective stimuli which are novel, and as a consequence also reduces fear. It is assumed that the mother is a highly familiar stimulus complex.

Berlyne (1967) has subsumed novelty under the broader category of collative variables. These are stimulus conditions of the environment which have the capacity to induce conflict by their own incongruity or by their incongruous relation to existing cognitive structures of the organism. The induced conflict raises the level of arousal of the organism, thereby altering its affective state. Differing slightly from King, Berlyne proposes that arousal, and therefore novelty, is curvilinearly related to an affective continuum ranging from indifference through pleasure or reward to displeasure or noxiousness. Berlyne (1967) postulates that because of the higher activation threshold of the aversion system than of the reward system in the brain, moderate increases in arousal mobilize the reward system, while more intense increases in arousal cause the aversion system to become active and also to exert an inhibitory influence on the reward system. According to both King and Berlyne, high degrees of novelty would result in noxious levels of arousal. Also under Berlyne's theory the arousal by any stimulus condition (internal or external, chemical or mechanical) will be added to the arousing effects of novelty and jointly determine the affective state of the organism. Of course, which specific stimulus configurations have arousing consequences will depend upon the prior experience of the organism and the structure of its nervous system. Stimuli which are frequently or extensively attended to undergo reductions in their capacity to arouse, a consequence which may make them less aversive, more rewarding, or, if carried far enough, affectively indifferent. Thus, the proximity of an attached object, because of its high familiarity, may serve to reduce general arousal, perhaps moving it from noxious levels to rewarding or affectively indifferent levels.

Does the concept of attachment have application to friendships and social ties which do not have their origin in infancy, and will these social relationships have the same functional properties as attachments in infancy, namely, the inhibition of distress or the evocation of positive affective states in the subject by proximity to the object of attachment? This is the basic question to which this investigation was addressed. The empirical evidence bearing directly on this question is rather limited. However, a fair amount of experimental work with higher mammals and the theoretical models derived from it have some relevance.

A case in point is Cairn's (1966) stimulus-sampling theory of attachment. Cairns has attacked the notion of an early immutable critical period for the formation of social attachments. It is his view that many mammals will form attachments to any salient stimulus complex, animate or inanimate, to which they have had prolonged exposure. He theorized that, via exposure, the stimulus complex becomes a pervasive link in the stimulus chains supporting many behaviors in the animal's repertoire. Scott (1963) ascribes an equally important role to exposure. It is his view that the emergence of fear of the strange as a competing motive signals the end of a period of easy attachment formation in that avoidance responses by the subject conflict with achieving the conditions of exposure necessary for the formation of an attachment.

Scott (1967) differs with Cairns as to the basic learning processes which underly the formation of attachments. He sees the process in dogs as being rooted in two motives which appear to develop at different times in the maturation of the puppy. The first, appearing at three weeks of age, is absence of the familiar. Removing accustomed stimulation by putting the pup in a strange place or leaving it in a familiar place such as the den while removing the mother and litter mates results in distress vocalizations which average 140 per minute in beagle pups. The second motive, which appears to be fully developed by 12 weeks of age in puppies, is fear of the strange. Both of these noxious emotional states are relieved by returning to familiar places or contact with familiar individuals. It is the reduction of these noxious states which Scott conceives of as an internal reinforcement which establishes and supports attachment to familiar individuals. This view is congruent with Berlyne's arousal-reinforcement model. The plausibility of the operation of similar mechanisms in the development of specific attachments in humans is supported by the following findings of Schaffer and Emerson

(1964): fear of strangers and specific attachments were typically first manifested within one month of each other, and, in their sample of 60 infants, the age of onset and specific attachment and of fear of strangers were positively correlated.

If the reduction of noxious degrees of strangeness or novelty serves as the reinforcement for attachment behavior, or if proximity is maintained simply because the locus of the attached object is the most familiar stimulus context in the environment, to what age might the aversiveness of the unfamiliar be sufficiently noxious as to support attachment related behavior? The phenomenon of cultural shock seen in adults who go to foreign countries suggests that rather than the fear of the strange decreasing with age, the individual simply becomes more extensively familiar with his potential environment, and therefore infrequently experiences noxious degrees or strangeness. Bovard (1959) has argued that the presence of social stimuli reduces aversive responses and physiological reactions to stress. He has proposed a physiological model to account for the effect. On the other hand, Zajonc (1965), after examining the literature of social facilitation, concluded that social stimuli have an arousing or drive inducing effect which facilitates dominant responses. I propose that the presence of _unfamiliar_ social stimuli has an arousing and stress inducing effect on the subject, but that the presence of _familiar_ social stimuli has a stress reducing effect on subjects already highly aroused or under stress and a mildly arousing effect on unaroused subjects of any age.

By the age of four the child spends as much waking time with his peers as he spends with parents. When he starts to school the proportion of time with peers increases. Thus the conditions are provided for weakening attachments to parents and for strengthening attachments to peers. Harlow has postulated a natural time for the emergence of peer-peer affectional bonds. He suggests that the appearance of such bonds is universal in the primates (Harlow and Harlow, 1966). Recent work by Sackett (Sackett, et al., 1967; Sackett, 1966) suggests that the switch from a preference for maternal proximity to peer proximity may be augmented by maturation changes in the organism. Infants raised from birth to 8 months solely in contact with the natural mother preferred proximity to a peer in a choice situation where the mother and a strange peer were the alternatives, whereas they chose the mother when the alternatives were the mother and a strange adult female monkey (Sackett, et al., 1967). Further evidence of an innate basis for the switch to peer preference during late infancy in monkeys was provided by a study of activity induced in infant monkeys raised in total isolation from birth by projecting color slide pictures on a wall of their living quarter. Among a wide variety of subject matter of varying complexity, the infants were most activated by pictures of other infant monkeys. This effect emerged at 8 weeks of age and continued to the end of testing at 9 months of age. It was manifested by differential play, climbing, verbalization, and exploration of the picture (Sackett, 1966). If the saliency or emotionally arousing properties of the stimulus object influences the strength or probability of attachment formation as suggested by Cairns (1966) and Scott (1963), the possible existence of innate arousal tendencies to specific stimulus configuration may exert a powerful influence over which objects eventually become attached.

Empirical investigations with humans beyond the infancy period examining the hypothesis that attached or familiar social stimuli are more effective stress reducers than individuals with whom there has been no previous interaction are few in number but, for the most part, supportive of the hypothesis (Burlingham and Freud, 1943; Kissel, 1965; Mandelbaum 1952; Marshall, 1951, and Schwarz, 1968).

Only in the studies by Mandelbaum, Marshall, and Kissel could the familiar social stimulus be construed as a peer and, of these studies, only Kissel's was experimental. In Kissel's experiment (1965) stress was conceptualized in terms of activation level. The index of activation level employed was palmer skin conduction. College students were given four perceptual reasoning tasks to solve which were in fact insoluable. Skin conductance was continuously recorded while the S̲s worked in one of three social contexts, (a) alone, (b) in the presence of a stranger working on an unrelated task, and (c) in the presence of a close friend also working on an unrelated task. The analysis of skin conductance level indicated highest activation in the Alone Condition, and lowest activation in the Friend condition. Activation in the Friend Condition was significantly lower than in the Stranger and Alone Conditions. The latter did not differ significantly. Subjects high in test anxiety had significantly higher activation levels. The need affiliation level of subjects was found not to be significantly related to social stimuli as stress reducers. The results failed to support an extrapolation from Schacter's (1959) theory that stress arouses affiliative tendencies. Kissel theorized that gratification of affiliative needs would reduce stress. Since stress level was not a function of the strength of the affiliative motive disposition (as measured by the Atkinson method), and the presence of a stranger did not reduce stress, his hypothesis received no support. With similarity to the views of Schwarz (1968) and Bronson (1968) and in congruence with the arousal-reinforcement model of Berlyne (1967), Kissel concluded that social stimuli act directly in reducing stress and that the presence of a friend, an emotionally pleasing stimulus, elicits positive feelings which tend to compete with negative feelings which are stimulated by stress. Less compatible with Berlyne's model and more in line with traditional S-R theory, Kissel further reasoned that friends in general have attributes similar to those who in the past have been dispensers of security, affection, and protection, whereas strangers lacking such attributes, will not call forth positive responses which would compete with the negative responses elicited by stress and, therefore, stress will not be appreciably reduced.

The overall plan of the present investigation was, first, to locate pairs of four and five year old children who appeared to have established a mutual attachment, i.e., they appeared to prefer the proximity of one another and each was unhappy about the other's absence. Observations were made of children who gave evidence of such attachments in a novel environment under three conditions: (a) in the presence of the attached object (friend), (b) in the presence of a child who was attached to some other child and a stranger to the subject, and (c) alone. The novel environment, a relatively unfamiliar room, contained four stationary toys, two of which were novel and two of which were familiar. As indices of distress and security, ratings were made of comfort, motility, and, in the case of the paired subjects of the Friend and Stranger Conditions, of verbal communication. It was hypothesized that proximity of an attached peer would have a distress inhibiting or security inducing effect which was greater than that which would be afforded by a strange peer. Therefore, it was predicted that children in the Friend Condition would show greater comfort, greater motility, and more verbal communication. It was also predicted that children in the Friend Condition would score higher on the indices of security than children in the Alone Condition, because of the distress inhibiting effect of proximity to an attached peer. No predictions were made as to potential differences between the Stranger Condition and the Alone Condition, since strong arguments can be made for differences in either direction. For example, it could be argued from

Berlyne's arousal-reinforcement model, that the novelty of the stranger would be arousing and added to the novel cues of the room yield a total arousal level which would have negative affective consequences. On the other hand, if the room itself were not sufficiently arousing, novel cues from the stranger might not by themselves be strong enough to exceed the aversion threshold. From a different tack one might argue that since even a strange peer is on a generalization continuum with the attached peer, in this case being of the same age, sex, and race as the attached peer, the strange peer's novelty may be slight and rapidly dissipated, yielding a functional equivalence to an attached peer. In the latter case, the Stranger Condition would result in greater security and distress inhibition that the Alone Condition. One might also argue that being alone is more novel for a nursery school child than being with a strange peer, since a variety of peers are his daily companions and he is seldom alone.

Concerning the toys, it was predicted that children in the Friend Condition would spend more time playing with the novel toys than children in the Stranger and Alone Conditions, and less time playing with the familiar toys than children in the Stranger and Alone Conditions. The rationale for this prediction was that the security inducing effect of proximity to the attached peer would keep the arousal induced by exposure to the novel toys below the aversion threshold, wheras exposure to further novelty in the absence of the familiar cues of an attached peer would be more likely to exceed the aversion threshold postulated by Berlyne (1967), and motivate avoidance of further voluntary exposure to high levels of novelty.

## Method

The Ss were drawn from the total population of 96 children at the Liverpool Laboratory Nursery School in Liverpool, New York. The school had three morning and three afternoon classes, each with 16 children. In each class the head teacher and the assistant teacher were asked to list, independently of each others' judgement, the four best classroom friends of each child in the room, and then to rate the degree of closeness of each friendship on the following scale.

A. Often likes to be near this child. If given a choice of partners, would frequently choose this child. Would miss this child if he (she) were absent from school. Would like to spend much of the class day near this child.

B. Frequently spends part of the class day near this child. If given a choice of partners, would occasionally choose this child. May notice this child's absence from school.

C. Doesn't mind being near this child, but would rarely seek him out. Occasionally plays near this child. Would not notice this child's absence from school.

Examples were provided and it was pointed out that all of a child's friends might be of the same degree of closeness and that some child may have no friends or only those of the least degree of friendship. These ratings of friendship were used as a basis for identifying like-sex mutual friendship pairs among the children in each class. The minimum criterion permitting two children to come under consideration as a friendship pair was that each child be listed by both teachers as a friend of the other child. In addition, the degree of friendship each child held for the other must have been rated B or higher on the above scale by one of the two raters. Using this criterion, 30 friendship pairs were identified.

Experimental situation. Observations were made in a 14 x 15 ft. room of one of a suite or rooms used for testing, located in the same building as the classrooms but in a separate corridor not frequented by the children. The experience of the children with this part of the school was restricted to those occasions when they were escorted there for purposes of testing. The experimental room was unfamiliar to the children in the sense that it was not a part of their everyday experience and no child had had free play experience there, although most Ss had been in the room on a prior occasion. The experimental room had no outside windows. At each end of the 15 ft. dimension was a one-way mirror 4 x 8 ft. which gave observers in the adjacent room full view of activities in the experimental room. The walls of the experimental room were cream colored and the ceiling, which held recessed fluorescent lights, was white. The floor was covered with a burnt orange and black tween carpet. For this experiment, and differing from any S's prior exposure to it, the floor area was marked off with 1 in. masking tape in a grid configuration with 2 ft. squares. Independent of the grid, the floor was criss-crossed with 2 in. masking tape from corner to corner, thus dividing it into four quandrants.

Familiar toys. Each triangular quandrant of the room contained a toy which was situated close to the wall and centered on the side of the quandrant bordering the wall. On each 14 ft. side was a familiar toy. One of these was a black flannel board (Instructo) 3 ft. x 2 ft. which held 15 felt symbols of various colors and shapes ranging from 1 in. to 2½ in. across. The flannel board which was supported by a metal stand, faced the center of the room and was tilted back slightly with the 2 ft. side against the floor. This position permitted a 4 year old child to play comfortably in a standing or seated position. Each of the classrooms contained an identical flannel board to which the children had free access. The second familiar toy consisted of a steering wheel and gear shift mounted on a post secured to a stable base. The base also held two spring loaded pedals. Three sturdy wooden blocks were stacked in front of the steering wheel to serve as a seat. This toy was situated in the center of the wall opposite the flannel board. A child seated at the steering wheel would have had his back to the flannel board. The steering wheel and flannel board were selected from among the toys in the Nursery School because of their reported popularity with children of both sexes and because they were less likely to be moved about the experimental room than other potential selections. The steering wheel was the only one in the nursery; however, because of its popularity, it had been rotated religiously from one classroom to another at weekly intervals. Each child had had approximately 5 weeks of prior access to it.

Novel toys. In the selection of novel toys an attempt was made to obtain toys which had the following characteristics: likely to remain where stationed by the experimenter; low probability of familiarity to any child; complex in appearance; and capable of producing some unexpected stimulus change when played with. No commercially available toy completely satisfied all of these requirements; therefore, two toys were constructed from a combination of toys. One of these (Copter) had as its basic element a battery operated field helicopter toy (Sears, Roebuck and Company). It consisted of an orange plastic box 12 x 10 x 5 in. designed to be worn on the back with straps. The front of the box contained numerous dials and buttons. A 14 in. drive shaft extending from the top supported a 23 in. plastic rotory blade. The helicopter toy was affixed to a black pedistal consisting of a hollow plywood post, 23 in. high secured to a 34 x 20 in. plywood base for stability. Pulling down a lever on the right hand side of the toy caused a clicking noise, rotation of the rotor, and the alternate flashing of a red and a blue light on the coptor body. A further depression of the lever doubled the tempo of this action. The battery operated sound producing mechanism of a toy rifle (Sound-O-Power M-16 Military Rifle, Louis Marx & Co., Inc) was installed in the hollow pedistal. A lever trigger was prominantly situated on the control panel of the copter. Each pull of the trigger produced a discrete burst of recorded gun fire and battle sound from the sound mechanism in the pedistal.

The major element of the other novel toy (Dash) was a yellow plastic play dashboard (Firebird 99 Sports Car, REMCO Industries, Inc.) As with the copter, the dashboard was attached to the top of a black pedistal consisting of a 17 x 21 in. base and a hollow post 18 in. high, and 10 x 4 in. in cross section. The dashboard, which was 12 x 10 x 5 in., contained a clear plastic windshield with a wiper operated from a knob on the dash; a 7 in. steering wheel which flashed a right or left turn signal upon right or left rotation; an air horn on the steering wheel; and a small glove compartment. A battery operated device which made a variable intensity motor sound was attached inside the pedistal (V-RROOM Hot-Rodder Engine, Mattel, Inc.). The potentiometer switch was affixed to the left of the steering wheel. Clockwise rotation of the key produced first a brief loud sound immediately followed by a softer and then gradually increasing sound intensity. Also attached to the pedistal below and adjacent to the dash was a yellow plastic Surprise Box (Kohner Bros., Inc.) consisting of spring loaded compartments, each of which could be opened by manipulating its independent and unique mechanical latch. The smiling faces originally situated in each compartment were replaced with color pictures of animals and people in threatening poses.

Design and Procedure. The thirty mutual friendship pairs formed on the bases of teachers' ratings of friendship were randomly assigned to three experimental conditions, the Friend, Stranger, and Alone Conditions. The purpose of the pairing of Ss before random assignment was to prevent the introduction of a personality bias which would have arisen if an attempt was made to utilize the entire Nursery School population by placing children without friends in the Stranger or Alone Conditions. Subjects assigned to the Alone Condition were observed individually in the testing room. Those assigned to the Friend Condition were observed jointly in mutual friendship pairs. The Ss assigned to the Stranger Condition were not observed in mutual friendship pairs, rather a new pairing of these Ss was made. To insure a status of stranger and nonfriendship, restrictions were placed on this reassignment to pairs: members of a new pair had to be from different classrooms, and they could not ride to school in the same car pool. Also, the requirement of like-sex pairing was retained for the Stranger Condition, and for practical reasons, although from different classes, the classes had to meet in the same session, i.e., morning or afternoon.

The experimental procedure was identical for Ss observed individually in the Alone Condition and for those observed in pairs in the Friend and Stranger Conditions. The E went to the classroom and the teacher called the S (s) requested to come to the doorway where they were introduced to E, a female graduate student. The E remarked as follows: "We are going to go to one of those rooms down the hall from Miss Robinson's office (the school director). There I have some pictures and I am asking the boys and girls in your school to look at these pictures and to tell me which ones they like and which ones they don't like." In the case of strange pairs, this explanation was given to the first child picked up and then repeated when the second child was picked up from another classroom. Children who were strangers were not introduced but, as with the children in the other conditions, were taken directly to the experimental room.

The experimenter walked the child or children to the center of the experimental room and remarked as follows:

> "Oh my! I forgot to bring in my pictures. They are still out in my car and it is on the far side of the parking lot. I am sorry but I will have to go get them. That's all right because, while I'm gone, you will have a chance to play with the toys in this room. You see, there is a toy over there and one over here and one here and one over

here. (These instructions were given while turning the S(s) to face the toy at each side of the room). You may play with all of them while I am gone. It will take me about 5 minutes, but as soon as I get the pictures from my car I will come back here to this room and have you look at them."

At this point E departed and at the conclusion of the 5 min. observation period reentered the room with a pack of pictures in hand. After spending a few minutes eliciting each child's reactions to the pictures, E returned the children to their classroom. The experimental instructions seemed to avert any strong fear of abandonment on the part of the Ss since no child cried during the observation period.

Live observations and ratings. A videotape recording was made of the activities of the children in the experimental room. To obtain a picture of the entire floor area, a television camera equipped with a wide angle (10 mm. focal length) lense was mounted on the wall pointing directly into a convex mirror, which reflected the entire floor area. The mirror, a 24 x 15 in. Detecto mirror, was suspended at a slight angle toward the room at the junction of the ceiling and the wall directly above the steering wheel toy. Sound was picked up from a microphone mounted on the opposite wall. Most of the ratings were made from this videotaped record.

Two ratings were made life: One of looking behavior, and the other of overall comfort. Two raters were present for each; the E who brought the children to the experimental room and another graduate student who started the videotape equipment. The latter was not informed as to the friendship status of paired Ss. For paired Ss a record was made with an Esterline-Angus event recorder of the duration time and frequency of the occasions that each child looked at the other child in the room. Each rater observed one of the Ss and depressed a double pole switch for the duration of each event. The event recorder and switches were wired in such a way that a third channel operated on those occasions when the Ss were simultaneously looking at each other. No rating of looking behavior was made in the Alone Condition. Immediately after each session global rating of the apparent comfort were made. The following 5-point scale was employed for this purpose:

0 - Crying

1 - Despondent or frightened facial expression but no crying. Little or no play with toys. Little or no conversation or comment.

2 - Watchful but not despondent or frightened facial expression. Cautious, timid, or somewhat hesitant play with toys, especially at first. Few comments or verbalizations.

3 - Relaxed facial expression. Active play with toys during most of the session but possibly some hesitation initially. Some conversation and comment.

4 - Smiling, cheery facial expression. Active play with toys with no hesitation. Frequent animated conversation and comment.

All Ss were rated independently by both raters. The product-moment correlation between these ratings was .88 (df = 55). Since these two raters knew the general hypothesis under investigation and one knew the friendship status of pairs, ratings of comfort were also made from the videotape recordings by two new graduate students who had no knowledge of the research project. The comfort ratings of these two raters correlated .71 (df = 55). The average of the two ratings made live correlated .69 with the average of the two ratings made from videotape. It should be noted that because of the wide floor area covered and the overhead angle of the TV camera, facial expressions were not discernable of the videotape recordings and hence could not serve as a determinant of comfort ratings made from videotape.

Ratings from Videotape recordings. Two raters unfamiliar with the hypotheses were involved in rating each variable. Before rating a variable, a pilot tape of a pair of Ss was replayed as often as required to familiarize the raters with the procedure. When paired Ss appeared on the videotape, the raters decided beforehand who would rate which S. All Ss were rated on a given variable before the tapes were replayed to rate a different variable. To check for possible bias due to communication between raters during rating in the Alone Condition, all variables were rerated for Ss in the Alone Condition by a third rater, also unfamiliar with the hypotheses. The ratings of raters A and B, the initial raters, and of rater C were intercorrelated to obtain estimates of interrater reliability for each variable. For the 17 Ss in the Alone Condition the lowest of these interrater reliability coefficients was .91. Since half of the Ss in the Friend and Stranger Conditions were rated by rater A and the other half by rater B, the rating of raters A and B were averaged for Ss in the Alone Condition in order to balance any systematic difference in the rating styles of the two raters.

One variable, fine movement, an index of activity or motility, consisted of a frequency count of lines touched or crossed by S during the 5 min. observation period. The lines in question were those of 1 in. masking tape which divided the floor into a grid configuration of 2 ft. squares. A second motility variable, gross movement, consisted of a frequency count of the total number of quandrant entries or changes during the 5 min. observation period. In this case an entry was not counted unless the S spent 2 sec. or more in the quandrant, since the S was likely to step momentarily into an adjacent quandrant while moving from the quandrant at one end of the room to the opposite quandrant.

Another set of related variables was the percentage of the total 5 min. observation period spent in each quandrant. These scores reflected the relative attractiveness of the toys located in each quadrant; therefore, the scores were named for the toy in the quadrant. The time record was made with the event recorder referred to above. For this task it was equipped with two multiswitch frames, each containing a separate button for each quadrant. When two raters viewed a videotape of a pair of Ss, they began by depressing the button assigned to the quadrant in which their S was situated. When Ss changed quadrants, the button associated with the new quadrant was depressed, automatically restoring all other buttons. Each button independendently operated a channel of the event recorder, and was also wired in series with the button for the same quadrant on the other multiswitch frame. This circuitry produced an additional record of the frequency and duration of simultaneous occupancy of each quadrant by both Ss. As an index of the degree of proximity maintained by paired Ss, a conjoint presence score was calculated for each pair. It consisted of the percentage of the total time that paired Ss occupied the same quadrant and, thus, were less than 7 ft. apart. The gross movement score referred to above was also obtained from the event recorder record by simply summing the frequency with which each quadrant was entered.

Three scores were derived primarily from the audio information on the videotape recordings. One was the total frequency with which the gun on the Copter was triggered by each S, another was the total time (in sec.) that the V-RRoom motor on the dash was run, and the third was the total time spent in communicative verbalization by each S. Subjects in the Alone Condition were not rated for verbalization since so little occurred. All Ss in the Friend and Stranger Conditions were rated independently by two raters. The resulting sets of scores correlated .94 (df = 38). The average of the two ratings was employed in statistical analyses.

IQs from the Stanford-Binet Intelligence Test (1960 Revision, form L-M) were available for most Ss from testing done within six months of the experiment.

## Results

Table 1 presents mean scores on four variables for males, females, and for the sexes combined in the Friend, Stranger, and Alone Conditions. The variables are comfort in the experimental situation rated live and rated from videotape recordings, fine movement (crossing lines of 2 ft. squares), and gross movement (crossing lines marking quadrants of the floor). An analysis of variance (3 x 2 factorial design, conditions X sex, method of unweighted means) was applied to all four variables. These analyses were followed in each case by planned orthogonal comparisons of the Friend and Stranger Conditions and of the Alone Conditions with the combined Friend and Stranger Conditions. The remaining possible comparisons between individual means, i.e., Friend vs. Alone, and Stranger vs. Alone, were made by the Newman-Keuls procedure (Winer, 1962, pp. 80-85 and 101-103.) The Newman-Kuels test maintains appropriate alpha levels for multiple a posteriori comparisons. The same method of statistical analysis was used for the other dependent measures reported below.

Insert Tables 1 & 2 about here

It can be seen in Table 2 that the main effect of conditions was significant in the analysis of both ratings of comfort. With the live rating of comfort the effect of sex approached significance (p .10). Females tended to be more comfortable than males. No sex difference was suggested by the videotape ratings of comfort. As with all other dependent variables, the interaction between sex and conditions was not significant with either rating or comfort. Since the interaction between conditions and sex was not significant, the sexes were combined in comparisons of the simple effects of the three treatment conditions. As shown in Table 2, both ratings of comfort indicate that Ss in the Friend Conditions were significantly more comfortable than Ss in the Stranger Condition. Thus, the hypothesis that the proximity of an attached individual facilitates security and positive affective states is supported by this finding. There are some curious descrepencies between the results obtained with the live ratings of comfort and the videotape ratings. With the live ratings the Friend and the Stranger Conditions did not differ in comfort, and Ss in the Alone Condition were more comfortable than Ss in the Stranger Condition. These differences suggest a discomforting influence of the proximity of the stranger. However, the videotape ratings of comfort yielded the contradictory finding that the Stranger and Alone Conditions elicit a similar degree of comfort, a degree which is significantly lower than that for the Friend Condition. Thus, the videotape ratings of comfort, which were based on a more restricted sample of behavior but which also eliminated the possibility of rater bias, yielded results which supported the hypothesis of greater security in the presence of an attached individual.

The analysis of variance of fine movement scores (Tables 1 and 2) indicated that the main effect of conditions was significant beyond the .01 level. Male Ss tended to have higher fine movement scores but the main effect of sex did not reach significance (p .10). There having been no interaction of conditions and sex, male and female Ss were combined in comparison of the simple effects of the three treatment conditions. It can be seen in Table 2 that, as predicted, the mean of fine movement in the Friend Condition was significantly greater than the means in both the Stranger Condition and the Alone Condition, and that the Stranger and Alone Conditions did not differ significantly. The order of magnitude of the means for gross movement paralleled the order for fine movement (Friend Stranger Alone). Although the main effect of conditions was not significant, the planned comparison of the combined Friend and Stranger Groups differed significantly from the Alone Group. The Newman-Kuels comparison of the Friend and Alone Conditions was also significant. Both movement scores yielded results which are supportive of the hypotheses of greater security in the presence of an attached individual. The results with the fine movement measure were more strongly supportive of this hypothesis than those with the gross measure of movement.

Insert Table 3 about here

### Attraction to Novel and Familiar Toys

A summary of the data and the analyses of variance relating to the attractiveness of the four toys for male and female Ss in the three experimental conditions are presented in Table 3. The basic score employed here was the percentage of the total time that a child spent in the quadrant containing the toy indicated. It can be seen that the Dash, a novel toy, was by far the most attractive toy for all Ss. On the average, Ss spend 69.0 percent of the total time in the quadrant containing the dash. The other novel toy, the Copter, was the second most attractive toy, with Ss spending 23.0 percent of the time in its quadrant. Much lower in relative attractiveness were the two familiar toys, the Steering Wheel and the Flannel Board. Subjects on the average spent 5.3 and 2.7 per cent of the total time, respectively, in the quadrants containing those toys. Despite the fact that the familiar toys were selected, in part, for their popularity in the classrooms, their attractiveness paled in comparison with that of the novel toys, irrespective of sex or the experimental condition of the Ss.

Differences between the three experimental conditions in attraction to each toy were small in comparison with the differences in overall attractiveness of the four toys. Contrary to prediction, the Friend and Stranger Conditions did not differ in attraction to any toy. As can be seen in the bottom half of Table 3, the analysis of variance of the percentage of total time spent in the Copter quadrant yielded a significant main effect for conditions and sex, whereas there were no significant effects in the comparable analyses of variance for the other three quadrants. The planned comparison of the mean of the Friend Condition with the mean of the Stranger Condition for the Copter quadrant was not significant (F 1). Newman-Kuels comparisons of Friend vs. Alone and Stranger vs. Alone were both significant at the .05 level, and both in a direction indicating less attraction to the Copter in the Alone Condition. With regard to the percentage occupancy of the Copter quadrant, as well as Flannel Board quadrant and Steering Wheel quadrant, the distributions tended to be J-shaped and positively skewed. Also the means of cells tended to be proportional to the standard deviations, as can be seen in Table 3. As a precautionary measure, in these three instances, the analyses were recomputed with the following transformation: Log (score +1). This transformation eliminated the heterogeneity of variability, but had only slight effect upon the significance levels. The significant sex effect for the Copter quadrant dropped from .05 to .10, while the conditions effect remained significant at the .05 level, and the Newman-Kuels comparison of Stranger vs. Alone dropped below the .05 level of significance. The 2 x 3 analysis of variance for Dash, Flannel Board, and Steering Wheel yielded no significant effects. Therefore, contrary to prediction, the Friend and Stranger Conditions did not differ in attraction to any of the four toys.

With regard to the predicted differential reaction of the three groups to the familiar as opposed to the novel toys, it may be pointed out that the relative magnitudes of means in some cases conform to prior predictions relative to greater insecurity in the Alone Condition, i.e., for the Copter quadrant and for the Flannel Board Quadrant; however, the differences are opposite to prediction in the Dash quadrant and the Steering Wheel quadrant.

Both the Dash and the Copter were equipped with devices which made rather loud sounds. These devices were under the voluntary control of the *S*s. Of the total sample 88 per cent turned on the Vrroom motor or the Dash and 68 per cent fired the gun on the Copter. Some *S*s sought a continuation or repetition of these auditory stimuli while others seemed to avoid further stimulation. Chi-square comparisons of the numbers of *S*s included in the cells of Table 4 with those of Table 1 indicates that the proportions of *S*s who activated these devices at least once did not differ appreciably across experimental conditions. (Motor $x^2$=3.62, df=2, p .05, Gun $x^2$=1.15, df=2, p .05).

Insert Table 4 about here

The analysis of variance of the duration of motor operation data summarized in Table 4 indicated no differences as a function of sex or experimental condition. (The *F*s for both main effects and the interaction were less than 1.0). A comparable analysis of the frequency of gun firing yielded a significant main effect for conditions (F=3.89, df=2/33, p .05) and a significant main effect for sex (*F* = 10.15, *df* = 1/33, *p* .01). The interaction was not significant (*F* 1). It can be seen in Table 4 that boys fired the gun with more than triple the frequency of girls. In planned comparison tests the Friend and the Stranger Conditions did not differ in frequency of gun firing (*F* 1), but these two conditions combined did differ from the Alone Condition (*F* = 6.25, *df* = 1/33, p .05). The frequency of gun firing for those who fired at least once was much less in the Alone Condition than in the Friend and Stranger Conditions. It may be noted that the standard deviations of the Friend and Stranger Conditions are much larger than the standard deviation of the Alone Condition. A log transformation was applied to the gun shooting scores. It succeeded in equalizing the variances of the three groups but did not change significance levels from those reported above.

### Further comparisons of the Friend and Stranger Conditions

Four behavioral measures were obtained only on *S*s in the Friend and Stranger Conditions. Each of these measures reflected an aspect of the social interaction between the paired *S*s in these two conditions. Therefore, no comparable observations were possible for *S*s in the Alone Condition. Two of these measures, verbalizations and looking at partner, were rated separately for each individual. The other two were rated for pairs of *S*s in each condition. The means on all four variables are presented in Table 5 by condition and sex and by condition with males and females combined. Each variable was analyzed in a 2 x 2 analysis of variance design (conditions X sex, method of unweighted means). The analysis of the verbal communication measure produced a significant main effect for conditions (*F* = 8.43, *df* = 1/36, p .01). As anticipated, individual *S*s in the Friend Condition had a greater mean number of seconds of communicative verbalization than individual *S*s in the Strange Condition. There was a tendency for males to communicate more than females, especially in the Stranger Condition; however, neither the main effect for sex nor the interaction of sex and condition were significant (*F* = 2.26, *df* = 1/36, p .05, and *F* = .69, *df* = 1/36, p .05, respectively). The analysis of the number of seconds spent looking at the partner yielded no significant effects for conditions, sex, or the interaction of sex by condition (*Fs* 1). The looking behavior records of two *Ss* in the Friend Condition were lost due to a failure of the event recorder. Although the Friend and Stranger Conditions did not differ in mean looking time, the variance of the looking times of *Ss* in the Stranger Condition was greater than the variance for the Friend Condition (*F* = 2.78, *df* = 19/17, *p* .05, two-tailed). This finding suggests that being paired with a stranger may result in either increased looking or decreased looking at the stranger, relative to the time spent looking at a friend.

Insert Table 5 about here

It may be recalled that the conjoint looking score is the total number of seconds that a pair of *Ss* simultaneously looked at each other or returned one anothers gaze. It was anticipated that strangers would tend to avoid meeting one anothers gaze. The conjoint looking scores presented in Table 5 failed to support this hypothesis; in fact, the mean differences were in the opposite direction, i.e., the pairs of strangers had a longer duration of conjoint looking than the pairs of friends. Neither the main effects nor the interaction approached significance in the analysis of variance of conjoint looking scores. It may be noted that the average duration of this behavior for all pairs was only 10.0 sec., whereas, the maximum possible score was 300 sec. With this latter figure in mind, it is apparent that both conjoint and individual looking at the partner occupied only a small proportion of the *Ss*' time in the experimental situation.

With regard to the score of conjoint presence in the same quadrant, it was predicted that friends would spend more time together in the same quadrant than strangers because of the greater tolerance for proximity between attached individual than between strangers. The analysis of variance of the conjoint presence scores, for which group means are presented in Table 5, revealed no significant differences as a function of condition, sex, or the interaction of sex and condition; therefore, the above hypothesis was not supported. It was surprising to find that even strangers spent 46 per cent of the total time jointly occupying the same quadrant or ¼ of the floor space in the experimental room.

Insert Table 6 about here

A matrix of the intercorrelations of all the individual subject variables is presented in Table 6. Since it was hypothesized that comfort ratings, motility, verbalization, and the strength of the preference for novel toys (Dash and Copter) would all reflect security, one would expect significant positive correlations between these variables. Comfort, movement, and verbalization should be negatively correlated with time in quandrants containing familiar toys. Playing with the noise making aspects of the novel toys (Motor and Gun) was also thought to reflect greater security. In general, it appears that the matrix of intercorrelations provides qualified support for these anticipated relationships. A major discrepency centers about time in the Dash quadrant. It can be seen that movement is negatively correlated with time in the Dash quadrant. Also, it will be recalled that Alone Condition S̲s tended to have higher mean scores for Dash than did the other groups. Time spent in the Copter quadrant, on the other hand, fits very well the anticipated relationships. It is significantly positively correlated with both comfort ratings and with motility ratings. Also, occupancy of the Copter quadrant is strongly associated with Gun shooting. The failure of Copter to correlate positively with verbalization is contrary to the postulated security pattern. Verbalization has significant positive correlations with motility and positive correlations approaching significance with comfort. Thus, the matrix provides some support for the utility of the construct of security as an explanitory concept in the theoretical account of the behavior observed in this experiment.

## Discussion

The results from the two comfort measures, the fine movement score and the verbalization score provide clear support for the hypothesis that the proximity of an attached peer has a distress inhibiting or security inducing effect on children which is greater than that afforded by proximity to a strange peer. The level of comfort exibited by friends was quite high for both live and videotape ratings. It would appear than that not only was their an absence of distress, but that the children in the Friend Condition experienced strong positive effect. Construing this result within the Berlyne arousal-reinforcement model, one might infer, then, that the proximity of a friend in this novel and stimulated environment resulted in levels of arousal which were above the reward system threshold but not so high as to activate the aversion system.

Some curious differences were observed between the live rating of comfort and the videotape rating. The major discrepancy was between the magnitude of the comfort mean of the Alone Condition relative to the other conditions. In the live ratings the comfort of the alone _S_s was rated on a par with that of the Friend Condition whereas, with the videotaped ratings, it was rated as low as that in the Stranger Condition. It will be recalled that facial expression could not serve as a determinant of comfort ratings made from the videotapes. It is possible, then, that the discrepant results are due to the differential influence of this behavior on the scores. Since facial expression might be thought of as the best single index of affective state available in this situation, one might then speculate that *S*s in the Alone Condition were experiencing a more pleasurable affective state than *S*s in the Stranger Condition. One problem with this conclusion is that it is not consistent with the high correlation and greater correspondence across conditions between the videotape comfort scores and fine movement scores than between the live comfort scores and the fine movement scores. The relative immobility of the alone *S*s would suggest discomfort and a general inhibition of motility. On the other hand motility, although conforming to prior predictions as an index of security and distress inhibition, may not be highly correlated with affective state. It is likely that the videotape raters responded more strongly to the cue of motility in judging comfort than did the live raters of comfort who had the cue of facial expression available.

If the alone *S*s did in fact experience greater pleasure in the situation than *S*s in the Stranger Condition, it would suggest that the proximity of a stranger was more aversive than being alone in this situation. The variable of conjoint presence in the same quadrant also may have some bearing on the issue of the aversiveness of proximity to a stranger. It was anticipated that strangers would spend less time in the same quadrant than friends. This was not the case. The failure to find this difference is one of the most perplexing findings of the study. Two interpretations seem plausible: one is that strangers were not aversive. However, this is contradicted by both comfort ratings and fine movement scores. Another interpretation which seems more plausible is that the attraction of the novel toys was so strong that it completely over rode any aversion to the stranger. Its plausibility stems from the fact that the attraction was so strong that *S*s spent 92 per cent of the total time in the two quadrants containing the novel toys. Also the small percentage of time (12.6 that *S*s spent looking at one another provides further evidence that the toys were the major focus of attention. The greater variation in time spent looking at the partner in the Stranger Condition than in the Friend Condition indicates that some Ss in the Stranger Condition avoided looking at the stranger while others visually explored the stranger. These looking pattern differences suggest that the stranger was not affectively neutral and may have been mildly aversive.

In a review of stidies on social facilitation, Zajoc (1965) put forth a generalization which seemed to account for a large number of diverse findings which, on the surface, appeared mutually contradictory. He hypothesized that the presence of others as spectators or as coactors, enhances the emission of dominant responses. He further speculated that arousal, activation, or drive was the intervening state variable which accounted for the increased emission of dominant responses, which was manifested in facililitation of simple task responses and interference with complex task responses. He then proceeded to examine the evidence for the hypothesis that the presence of others increases the individual's general arousal or drive level. He found that much of the existing evidence was indirect and somewhat contradictory. In all of the social facilitation studies surveyed by Zajoc, it appears that the subjects and social stimuli employed were strangers to one another. In view of Berlyne's (1960) theory of the effects of novelty upon arousal, the generality of Zajoc's integrating explanatory hypothesis maybe limited to the effects of unfamiliar social stimuli. The results of this study with regard to the fine movement response in no way support Zajoc's hypothesis even when modified according to the Berlyne (1960) and King (1966) hypothesis of a linear relationship between novelty and arousal. This updated version of the Zajoc hypothesis would have predicted that Ss in the Friend Condition would have been less aroused, because of the proximity of the familiar peer, than Ss in the Stranger Condition who were in the presence of a strange (novel) peer. However, if we view locomotion as a dominant response, which seems reasonable, we find that the results are opposite to that prediction in that the Friend Condition had the highest motility and the Stranger Condition had significantly less motility as measured by the fine movement score. Also the Zajoc hypothesis, updated or not, would have predicted greater motility in the Stranger Condition than in the Alone Condition where no difference was found. Rather, these results conform perfectly to the prior prediction that increased security, induced by proximity to an attached peer, would facilitate motility. It would appear that response systems differ in the manner in which they are affected by arousal of varying degrees. Freezing or immobility is a response which sometimes occurs under conditions of fear. The later is probably a high state of arousal. On the other hand fear and threat are likely to produce flight. In this situation mild nonaversive degrees of arousal, which seem to have existed in the Friend Condition, (judging from the comfort ratings) seem to have had facilitory effects on motility. In the absence of any specific threat which would induce flight, the reduced motility in the Stranger Condition probably reflects the inhibitory influence of a higher and mildly aversive level of arousal.

Several of the studies cited by Zajoc could be interpreted in this way. Highly social animals were seen to be less active when alone. The absence of accustomed stimulation may have been highly arousing, aversive, and inhibitory, while the presence of accustomed social stimulation may have been rewarding, mildly arousing, and facilitory. This interpretation is supported by Kissel's observations of higher activation (skin conductance) in college students solving problems alone as compared with a condition of proximity to a close friend.

The results pertaining to percentage of time in the Copter quadrant and repeated gun firing, on the surface, seem to reflect the importance of another influence operative in the Friend and Stranger Conditions, that of imitation. Allport (1924, p. 274) theorized that "the sights and sounds of others doing the same things" augment overt motor responses. It may be that the higher rates of repeated gun firing in these conditions than in the Alone Condition were a result of this type of imitative facilitation. However, play with the Dash and running the motor attached there also seemed to require no such facilitation since the rates were as high in the Alone Condition as in the other two conditions. Furthermore, the probability of a subject in the Alone Condition locating and firing the gun at least once did not differ from the paired subject conditions. Rather it appears that this is an effect which is due to the aversiveness of the gun sound and the Copter in general for Ss in the Alone Condition. The variance difference between the Alone Condition and the paired conditions indicates that there were no subjects in the Alone Condition who had high rates of gun firing as there were in the Friend and Stranger Conditions. Although there are no objective standards which can be applied, in the judgement of the author, the Copter and the gun sound of the motor on the Dash was both louder and more noxious than that of the gun. These facts and judgements support the view that subjects in the Alone Condition were least secure and had the lowest tolerance for highly novel stimulation.

In view of the foregoing discussion it must be concluded that the Dash unfortunately was not a highly novel stimulus. It did bear similarity to a real automobile dash board which these Ss encountered everyday when transported to the nursery school in car pools. Furthermore, its higher degree of realism than the Steering Wheel toy may account for its high attractiveness even in the Alone Condition. In addition to providing an optimal degree of novelty, it may have elicited a large store of gratifying fantasies of mastery in Ss of this age. The familiar toys were reacted to with indifference in this situation, a fact which was contradictory to the experimental hypothesis. One might assume that the Ss did not find the total situation so extremely arousing that it was necessary to seek the most familiar stimulus available in order to reduce arousal below noxious levels. One subject in the Alone Condition did, however, spend over two-thirds of the total time with the familiar flannel board.

It may be concluded that the results of this experiment can be readily construed within the framework of attachment theory and Berlyne's arousal-reinforcement model.

Table 1

Means and Standard Deviations of Ratings of Comfort and Motility in the Friend, Stranger, and Alone Conditions

| Variable | Statistic | Condition | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Friend | | Stranger | | Alone | | Friend | Stranger | Alone |
| | | Male | Female | Male | Female | Male | Female | Combined | Combined | Combined |
| | N | 8 | 12 | 8 | 12 | 6 | 12 | 20 | 20 | 17 |
| Comfort | M | 3.38 | 3.88 | 2.75 | 2.79 | 2.83 | 3.59 | 3.68 | 2.78 | 3.32 |
| (Live) | SD | .92 | .43 | .89 | 1.05 | .41 | .74 | .69 | .97 | .73 |
| Comfort | M | 3.81 | 4.00 | 3.75 | 3.21 | 3.33 | 3.45 | 3.92 | 3.42 | 3.41 |
| (Video) | SD | .26 | .00 | .27 | 1.12 | .52 | .52 | .18 | .91 | .51 |
| Fine | M | 78.25 | 62.58 | 48.38 | 41.00 | 45.00 | 27.31 | 68.85 | 43.95 | 36.56 |
| Movement | SD | 19.25 | 33.84 | 13.40 | 30.49 | 44.34 | 20.06 | 29.35 | 24.86 | 30.69 |
| Gross | M | 13.38 | 9.00 | 9.25 | 8.33 | 8.00 | 5.73 | 10.75 | 8.70 | 6.53 |
| Movement | SD | 5.58 | 4.20 | 3.28 | 6.77 | 7.54 | 3.90 | 5.15 | 5.55 | 5.34 |

## Table 2

### Summary of Analysis of Variance of Comfort and Motility Ratings and Comparisons of Friend, Stranger, And Alone Conditions

| Source | df | Variable | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Comfort (Live) | | Comfort (Video) | | Fine Movement | | Gross Movement | |
| | df | MS | F | MS | F | MS | F | MS | F |
| Condition (A) | 2 | 3.25 | 5.19** | 1.34 | 3.65* | 5660.6 | 7.12** | 83.51 | 2.98 |
| Sex (B) | 1 | 2.50 | 4.00 | .08 | .22 | 2459.7 | 3.10 | 84.87 | 3.03 |
| A X B | 2 | .58 | .93 | .72 | 1.97 | 132.8 | .17 | 13.51 | .48 |
| Error | 51 | .62 | | .37 | | 794.7 | | 28.00 | |
| Planned orthogonal comparisons | | | | | | | | | |
| Alone vs. Friend, Stran. | | | .37 | | 2.24 | | 7.40** | | 4.36* |
| Friend vs. Stranger | | | 12.94** | | 6.79* | | 7.80** | | 1.50 |
| A posteriori comparisons[a] | | | | | | | | | |
| Friend vs. Alone | | | ns | | * | | ** | | * |
| Stranger vs. Alone | | | * | | ns | | ns | | ns |

[a] Newman-Kuels test. * p .05 ** p .01

## Table 3

Means, Standard Deviations, and Analyses of Variance of the Percent of Time Spent in each Quadrant of the Observation Room

| Variable | | Condition | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Friend | | Stranger | | Alone | | Friend | Stranger | Alone |
| | | Male | Female | Male | Female | Male | Female | Combined | Combined | Combined |
| | | N = 8 | N = 12 | N = 8 | N = 12 | N = 6 | N = 11 | N = 20 | N = 20 | N = 17 |
| Dash | M | 55.40 | 69.54 | 62.29 | 68.32 | 79.44 | 78.38 | 63.88 | 65.91 | 78.76 |
| | SD | 20.19 | 23.44 | 25.34 | 29.16 | 22.17 | 21.14 | 22.78 | 27.16 | 20.81 |
| Copter | M | 35.87 | 24.25 | 35.76 | 20.24 | 16.74 | 9.47 | 28.90 | 26.45 | 12.03 |
| | SD | 20.01 | 23.56 | 25.73 | 20.69 | 18.42 | 7.27 | 22.43 | 23.51 | 12.32 |
| Flannel | M | .15 | 2.13 | .93 | 2.69 | .59 | 7.47 | 1.34 | 1.99 | 5.04 |
| Board | SD | .29 | 3.56 | 1.02 | 4.59 | 1.45 | 21.95 | 2.89 | 3.65 | 17.70 |
| Steering | M | 8.57 | 4.08 | 1.01 | 8.76 | 3.23 | 4.68 | 5.88 | 5.66 | 4.17 |
| Wheel | SD | 19.24 | 8.76 | 1.34 | 21.77 | 6.82 | 4.05 | 13.63 | 17.03 | 5.03 |

2

### Table 3--Continued

| Source | *df* | Analysis of Variance | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Variable | | | | | | | |
| | | Dash | | Copter | | Flannel | | Wheel | |
| | | *MS* | *F* | *MS* | *F* | *MS* | *F* | *MS* | *F* |
| Condition (A) | 2 | 1375.1 | 2.36 | 1524.2 | 3.80* | 40.76 | .40 | 25.45 | .14 |
| Sex (B) | 1 | 541.3 | .93 | 1757.4 | 4.38* | 166.87 | 1.63 | 32.93 | .19 |
| A X B | 2 | 257.1 | .44 | 75.9 | .19 | 37.42 | .37 | 166.65 | .94 |
| Error | 51 | 581.8 | | 401.5 | | 102.11 | | 177.57 | |

*p .05

Table 4

Means and Standard Deviations of Seconds of Motor Running and Frequency of Gun Trigger Pulls for Children Who Used these Toys

| Variable | | Condition | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Friend | | Stranger | | Alone | | Friend | Stranger | Alone |
| | | Male | Female | Male | Female | Male | Female | Combined | Combined | Combined |
| | M | 66.71 | 72.90 | 77.86 | 63.78 | 65.92 | 79.18 | 70.35 | 69.94 | 74.50 |
| Motor | SD | 56.05 | 58.51 | 63.09 | 56.31 | 105.99 | 83.45 | 55.80 | 62.60 | 88.91 |
| Running | N | (7) | (10) | (7) | (9) | (6) | (11) | (17) | (16) | (17) |
| | M | 14.40 | 5.30 | 17.83 | 4.75 | 4.33 | 1.57 | 8.33 | 10.36 | 2.40 |
| Gun | SD | 14.05 | 4.47 | 12.48 | 4.10 | 3.51 | 0.79 | 9.43 | 10.68 | 9.12 |
| Shooting | N | (5) | (10) | (6) | (8) | (3) | (7) | (5) | (14) | (10) |

## Table 5

## Means and Standard Deviations of Variables Scored for Pairs of Subjects or Individual Subjects in the Friend and Stranger Conditions Only

| Variable | | Condition | | | | | |
|---|---|---|---|---|---|---|---|
| | | Friend | | Stranger | | Friend | Stranger |
| | | Male | Female | Male | Female | Combined | Combined |
| Individual | | | | | | | |
| Verbal-<br>ization | M | 27.25 | 18.12 | 18.12 | 7.88 | 25.48 | 11.98 |
| | SD | 17.24 | 18.43 | 5.85 | 7.56 | 17.56 | 8.50 |
| | N | (8) | (12) | (8) | *12) | (20) | (20) |
| Looking at partner | M | 39.67 | 34.67 | 44.50 | 35.92 | 36.33 | 39.35 |
| | SD | 12.86 | 18.74 | 30.92 | 26.64 | 16.78 | 27.96 |
| | N | (6) | (12) | (8) | (12) | (18) | (20) |
| Pair | | | | | | | |
| Conjoint looking | M | 11.00 | 6.00 | 13.25 | 11.67 | 7.67 | 12.30 |
| | SD | 6.08 | 3.22 | 9.43 | 7.58 | 4.69 | 7.89 |
| | N | (3) | (6) | (4) | (6) | (9) | (10) |
| Conjoint presence in a quadrant | M | 38.95 | 46.96 | 47.54 | 44.62 | 43.75 | 45.79 |
| | SD | 22.20 | 29.00 | 25.62 | 36.66 | 25.47 | 31.11 |
| | N | (4) | (6) | (4) | (6) | (10) | (10) |

## Table 6

### Intercorrelations of Ratings

| Variables | N | ComV | FMov | GMov | Dash | Copt | Flan | StWh | Gun | Mot | Verb | Look | IQ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Comfort (Live) | 57 | 69** | 13 | -03 | -12 | 29* | -15 | -13 | 14 | 24 | 19 | 03 | 13 |
| Comfort (Video) | 57 | | 26* | 18 | -19 | 30* | -19 | 00 | 28 | 43** | 29 | 06 | -01 |
| Fine Movement | 57 | | | 88** | -44** | 38** | -02 | 21 | 14 | -26 | 58** | 25 | -10 |
| Gross Movement | 57 | | | | -42** | 27* | 02 | 32* | 16 | -24 | 37* | 29 | -06 |
| Dash | 57 | | | | | -73** | -31* | -44** | -44** | 26 | -06 | -11 | -03 |
| Copter | 57 | | | | | | -17 | -15 | 70** | -13 | -09 | -01 | -04 |
| Flannel Board | 57 | | | | | | | 08 | -14 | -17 | 00 | -09 | 02 |
| Steering Wheel | 57 | | | | | | | | -23 | -23 | 04 | 20 | 10 |
| Gun (Scores > 0) | 39 | | | | | | | | | 20<br>(36) | -05<br>(29) | 19<br>(27) | -23<br>(38) |
| Motor (Scores  0) | 50 | | | | | | | | | | -15<br>(33) | -04<br>(31) | -25<br>(48) |
| Verbal Communication | 40 | | | | | | | | | | | 13<br>(38) | 12<br>(39) |
| Looking at Partner | 38 | | | | | | | | | | | | -16<br>(37) |
| IQ | 55 | | | | | | | | | | | | |

Note - Decimals omitted. Where N's for coefficients do not equal the smaller N of the two variables, the exact N is provided in parentheses.

*p  .05
**p  .01

## References

Ainsworth, Mary D. The development of infant-mother interaction among the Ganda. In B.M. Foss (Ed.) *Determinants of Infant Behaviour II*. New York: Wiley, 1963, pp. 67-112.

Ainsworth, Mary D.S. and Wittig, Barbara A. Attachment and exploration behavior on one-year olds in a strange situation. In B.M. Foss (Ed.) *Determinants of infant behaviour IV*. New York: Wiley, in press.

Berlyne, D.E. *Conflict, Arousal, and Curiosity*. New York: McGraw-Hill, 1960.

Berlyne, D.E. Arousal and reinforcement. In David Levine, (Ed.), *Nebraska symposium on motivation*, 1967. Lincoln: University of Nebraska Press, 1967.

Bovard, E.W. The effect of social stimuli on the response to stress. *Psychological Review*, 1959, 66, 267-277.

Bronson, G.W. The development of fear in man and other animals. *Child Development*, 1968, 39, 409-431.

Bronson, G.W. The fear of novelty. Psychological Bulletin, 1968, 69, 350-358.

Burlingham, D. and Freud, A. *Young children in war-time*. London: Allen and Unwin, 1943.

Cairns, R.B. Attachment behavior of mammals. *Psychological Review*, 1966, 73, 409-426.

Fraiberg, Selma. The origins of human bonds. Commentary, 1967, December, pp. 47-57.

Gray, P.H. Theory and evidence of imprinting in human infants. *Journal of Psychology*, 1958, 46, 155-166.

Harlow, H.F. and Zimmermann, R.R. Affectional responses in the infant monkey. *Science*, 1959, 1330, 421-432.

Harlow, M.K. & Harlow, H.F. Affection in primates. Discovery, 1966, January.

King, D.L. A theory of infant-mother relationship in mammals and birds. Psychological Bulletin. 1966, 65, 143-155.

Kissel, S. Stress-reducing properties of social stimuli. Journal of Personality and Social Psychology, 1965, 2, 378-384.

Mandelbaum, D.G. Soldier groups and Negro soldiers. Berkeley: University of California Press, 1952.

Marshall, S.L.A. Men against fire. Washington, D.C.: Combat Forces Press, 1951.

Morgan, G.A. and Ricciuti, H.N. Infants' responses to strangers during the first year. In B.M. Foss (Ed.) Determinants of infant behavior IV. New York: Wiley, in press.

Rheingold, Harriet L. The effect of a strange environment on the behavior of infants. In B.M. Foss (Ed.) Determinants of infant behavior IV. New York: Wiley, in press.

Sackett, G.P., Griffin, G.A., Pratt, C., Joslyn, W.C., and Ruppenthal, G. Mother-infant and adult female choice behavior in rhesus monkeys after various rearing experiences. Journal of Comparative and Physiological Psychology, 1967, 63, 376-381.

Sackett, G.P. Monkeys reared in isolation with pictures as visual input; evidence for an innate releasing mechanism. Science, 1966, 154, 1468-1473.

Schacter, S. The psychology of affiliation. Stanford: Stanford University Press, 1959.

Schaffer, H.R. and Callender, W.M. Psychological effects of hospitalization in infancy. Pediatrics, 1959, 24, 528-536.

Schaffer, H.R. and Emerson, Peggy E. The development of social attachment in infancy. Monographs of the Society for Research in Child Development, 1964, 29, (3, Whole No. 94).

Schwarz, J.C. Fear and attachment in young children. Merrill-Palmer Quarterly, 1968, October (in press).

Scott, J.P. The process of primary socialization in cainine and human infants. Monographs of the Society for Research in Child Development, 1963, 28, (Whole No. 85).

Scott, J.P. The development of social motivation. In D. Levine (Ed.) Nebraska symposium on motivating Lincoln: University of Nebraska Press, 1967.

# KINDERGARTEN "LEARNING TO LEARN" PROGRAM EVALUATION

Margaret Lay
Syracuse University

The purpose of the proposed study is to investigate the effects of alternative learning situations on young children's development. There is a sharp division of opinion among early childhood educators as to the manner in which children are most effectively educated which may be summarized as follows:

Many professional educators have considered optimal development to be, first, highly dependent upon naturational processes and, second, the by-product of spontaneous interaction of an emotionally healthy child in his environment. The teacher's responsibility, according to this view, is considered to be discharged by providing the "psychologically supportive" and "rich and varied" environment in which the child may initiate activities and thereby structure his own best learning situations.

Counter opinions have gained strength along with the current interest in the effects of early experience and intervention for the disadvantaged. Some of the programs claiming to have produced change in children's abilities have credited success to teacher-guided learning situations rather than the diverse opportunity for choice in what has been dubbed the "cafeteria" approach. Those questioning the effectiveness of the cafeteria programs doubt whether children, especially culturally deprived children, actually can identify for themselves from the global classroom environment the essential components of their optimal learning. They insist instead, that it is preferable for a teacher to structure sequential learning experiences in accordance with the individual learner's existing stage of development. The pro-"cafeteria" group, on the other hand, fear that children's creative abilities and capacities for self-direction will be thwarted if teacher-guided sequentially-arranged learning experiences are utilized with the young child. It is within this context of controversy that the proposed study is set.

This study will afford the opportunity to make comparisons for a range of subject behaviors between the sequentially-arranged teacher-guided programs and programs primarily focusing on children's self-initiated activities.

## REVIEW OF RELATED RESEARCH

Evaluations conducted by Van de Riet and Van de Riet (1966, 1967) of a sequential learning program for kindergarten children developed by Dr. Herbert Sprigle, Learning to Learn School, Jacksonville, Florida, have established it as a promising approach worthy of further evaluation. In the initial study three matched groups of culturally deprived Negro children were compared. An entire class group of 25 received the experimental program utilizing two classroom areas - a work-play area in which the total classroom group engaged in a variety of activities and a smaller room used for work with groups of four children in a sequential program of guided learning experiences. The children were taken from the regular classroom for daily short periods of participation in games and activities designed to "get the child to become active in the learning process". A second group received kindergarten training in an established "traditional kindergarten in the community, and a third group had no formal preschool training. Children in the experimental program were found to be significantly superior to either the "traditional kindergarten group or the no-treatment group on developmental measures including Stanford-Binet; Human Figure Drawings, Peabody Picture Vocabulary Test, Bender Gestalt test; Metropolitan Readiness Test; Visual decoding, Verbal encoding, Auditory-vocal association, Visual-motor association subtests of the Illinois Test of Psycholinguistic Abilities. The "traditional" trained group showed higher performance than the no-treatment group but the differences were not nearly as great as between the experimental and "traditional" groups.

The Sprigle program has also been experimentally utilized and evaluated for lower-middle class children and the following table showing scores from both studies reveals that the experimental program resulted in greater improvement for culturally-deprived children than for the lower-middle class children. At the termination of a nine-month program the culturally-deprived experimental children were functioning at levels similar to those of lower-middle class children exposed to "traditional kindergarten programs.

### Comparisons of Culturally Deprived and Lower-Middle Class Children at the Completion of Kindergarten

| Variables | | Learning to Learn Program | Traditional Program | No Program |
|---|---|---|---|---|
| | | Mean | Mean | Mean |
| Binet Intelligence Scores | Culturally Disadvantaged | 104.12 | 90.33 | 83.29 |
| | Lower-Middle Class | 112.83 | 107.33 | |
| Binet Vocabulary | Culturally Disadvantaged | 5.62 | 3.71 | 2.71 |
| | Lower-Middle Class | 7.00 | 6.19 | |
| Bender-Gestalt (error score) | Culturally Disadvantaged | 11.96 | 15.46 | 17.33 |
| | Lower-Middle Class | 7.91 | 11.48 | |
| Metropolitan Readiness Test - Total | Culturally Disadvantaged | 66.46 | 44.71 | 40.79 |
| | Lower-Middle Class | 66.78 | 47.38 | |
| School Readiness Screening Test | Culturally Disadvantaged | 20.08 | 13.79 | 13.21 |
| | Lower-Middle Class | 24.65 | 22.24 | |
| Sequin Form Board (time score) | Culturally Disadvantaged | 23.46 | 31.46 | 33.08 |
| | Lower-Middle Class | 21.26 | 21.62 | |
| Rail Walking (error score) | Culturally Disadvantaged | 10.92 | 31.83 | 28.21 |
| | Lower-Middle Class | 10.78 | 16.48 | |
| Human Figure Drawings | Culturally Disadvantaged | 16.33 | 10.04 | 7.08 |
| | Lower-Middle Class | 20.09 | 14.52 | |

Although the Sprigle program in toto has been demonstrated to make an impressive difference in children's development, especially culturally-deprived children, there remain unanswered many questions as to the factors most contributory to these gains. The highly-structured sequential tasks used in the experimental program were assumed to be the major contributing factors. Van de Riet, however, also noted the high level of "investment" and teaching competence of the author/director of the experimental program. The issue of teacher involvement must be accounted for by further evidence. A third possibility would also seem to warrant investigation. The evaluators did not determine to what degree the program of general classroom activities (beyond the special games and activities utilized as part of the guided-learning sequences) were different from those of the "traditional" classroom. Observers, including this writer, reported a richness of materials and opportunities for unstructured interaction and exploration quite distinctive from most "traditional" settings. The question, then, arises as to what degree the success of the program may be attributed to richness of general classroom program rather than to the small-group sequential learning experiences per se. The proposed study attempts to answer this question.

A fourth possibility, although seemingly less likely to account for developmental gains, is that simply being taken from the classroom in small groups and receiving "extra attention" is a significant experience in itself. Although Blank (1967) in a somewhat similar small-scale pilot study did not find gains on Stanford-Binet scores for two subjects receiving individual attention outside the classroom, other such comparisons have not been reported and cannot be dismissed as possibilities. The proposed study will provide a "placebo" situation to explore these effects.

## OBJECTIVE OF THE STUDY

The objective of the proposed study is to determine whether a teacher-guided, sequentially-arranged program of instruction for kindergarten children utilized in addition to a regular classroom program will be more effective in producing general intellectual gains and specified behavioral characteristics than two instructional alternatives. These alternatives are (1) participation in regular kindergarten programs, (2) participation in a special program of expressive activities in addition to the regular program. Assessments of subjects' performance on standard developmental measures and in a range of selected discrete situations will be obtained and utilized to determine differences between groups assigned to the above-mentioned alternative instructional situations.

Definitions:

For the experimental "teacher-guided sequentially-arranged program of instruction" children will be taken daily from the regular classroom in groups of three to six to participate in a pre-arranged sequence of activities under a teacher's direction. These sessions of approximately twenty minutes duration will be conducted in a room equipped with only those materials utilized from the day's instruction.

The placebo "program of expressive activities" will consist of approximately twenty minutes of daily participation by children in groups of three to five in activities such as easel painting, finger painting, puppet play, experimentation with musical instruments, clay work, etc. These sessions will also be conducted in a room equipped only with those materials utilized for the day's activity. Although the teacher in charge of these activities will provide materials and will use the children's utilization of these materials as a focus for converstional interchange, no effort will be made to provide sequential experience designed to produce particular accomplishments or products.

Subjects

The subjects of the study will be kindergarten children enrolled in an inner city Syracuse public school. Chidlren from six classes at Croton Elementary School will be utilized. Each of the classes is expected to have an enrollment of 20-25 children providing a pool of approximately 135 subjects. These will largely be Negro children and they are predominantly from economically and socially disadvantaged homes. A few of the children entering the kindergarten program will have had previous school experience in either the pre-kindergarten program of Syracuse City School District or in Summer Head Start. A few will have have been in Day Care programs.

Procedures

The manipulation of the program for experimental purposes will consist of providing additional experiences for some sub-groups beyond the base classroom experience. There will be three treatment conditions to be designated as Experimental, Placebo, and Control. Seven subjects from each class group will be randomly assigned to each of the respective conditions[1], totaling 42 for each across the six classroom groups.

Experimental subjects will be taken from the regular classroom for twenty minute daily periods to participate in a teacher-guided sequential-learning program. Placebo subjects will also be taken from the regular classroom for twenty minute daily periods but will participate in various expressive activities with a minimum of teacher direction and no pre-planned utilization of the materials to effect particular cognitive attainments.

---

[1] There is some question as to the best way to proceed for the formulation of these treatment sub-groups (N=6) within each group. It could certainly be argued that assignment by matching according to pertinent pre-test data is in order since the probability of getting skewed distributions in making random assignments into groups this size must be considered. However, after weighing the advantages accruing from random assignment, the following course is proposed. Children will be initially randomly assigned to sub-groups; however, if pre-test measures reveal significant differences, matching procedures will then be utilized for re-assignment prior to any experimental treatment.

Control subjects will have only the classroom experience.[2]

The twenty minute sessions for the experimental and placebo groups will be conducted by half-time teachers not involved in the regular classroom programs. Each of these teachers will be responsible for sequential learning programs for 21 children, 7 from each of three groups in session during the morning and the afternoon, and, similarly, for the placebo expressive activities for 21 children. The 21 experimental children for which each of these teachers is responsible will be re-grouped across class lines to more closely approximate desirable homogeneous instructional levels but only three to seven children will be in session for Experimental or Placebo activities at any one time. Similar re-groupings of placebo sub-groups will also be effected to maintain similarity of treatment.

The following diagram indicates the teacher assignment plans for the children enrolled in the six classes.

---

[2] It is recognized that the possibility of "ripple effects" between sub-groups within each class is a weakness of this design. However, the advantages of having three comparison groups of a common base experience in the classroom are considered, for this study, to outweigh the limitation of possible contamination. The effects would, at any rate, tend toward the elimination of experimental differences. If differences are found despite "rippling" possibilities there would be no reason to question the validity of those results.

| | Special Teacher 1 | | | | Special Teacher 2 | | | |
|---|---|---|---|---|---|---|---|---|
| Classroom Teacher A | C | Ex | Pl | | C | Ex | Pl | |
| Classroom Teacher B | C | Ex | Pl | | C | Ex | Pl | |
| Classroom Teacher C | C | Ex | Pl | | C | Ex | Pl | |

C - Control

Ex - Experimental

Pl - Placebo

The special teachers who will be conducting the experimental sequential learning program will receive training in the methods and materials of this approach under the direction of the principal investigator in consultation with Dr. Herbert Sprigle of the Learning to Learn School, Jacksonville, Florida during September, 1963. (See appendices for further descriptions of the experimental program.)

The function of the placebo program will be to provide a situation of similar duration, personnel, group size as the experimental condition but without effort to produce specific cognitive attainment through the utilization of sequential learning experiences. As indicated previously, the teachers will be instructed to use the products of the children's activities as a focus for conversation. The materials utilized in the placebo program will also be available in the regular classroom program and the way in which they are used probably will not differ to any significant degree. It is anticipated that the placebo situation will neither substantially add to nor subtract from the experience of the regular classroom program - although providing the same dimensions in regard to teacher-child ratio and physical setting as the experimental treatment.

As indicated in the diagram, special teachers are being assigned to both experimental and placebo teaching. The rationale behind this strategy is, of course, that the likelihood of teacher difference affecting results will be lessened if the same individuals are involved with the two treatment conditions. There exists the possibility, on the other hand, that teachers might prefer one of the approaches over the other and that these attitudes would influence teaching behavior. It is anticipated that this probability can be reduced through concerted efforts to emphasize to the Special Teachers that it is expected that significant contributions to children's development may accrue from both of the special programs. In other words, it will be reinterated to the special teachers throughout the study that both approaches are expected to make positive, although perhaps different, contributions.

To summarize, the proposed study will compare groups of kindergarten children who have been in varying treatment situations, designated as follows.

| | |
|---|---|
| Experimental - | Regular school program plus sequential learning program. |
| Placebo - | Regular school program plus expressive activities program. |
| Control I - | Regular school program only. |

## Evaluation

In evaluating children's progress at the end of the research period we plan to use both standard measures of development and data on subjects' behaviors in special task situations which seem especially relevant to the particular programs under consideration.

Standard measures to be administered include the Stanford-Binet (for which pre-scores will be obtained from all subjects in the fall) and the Peabody Picture Vocabulary Test. (For the latter, raw scores will be used for analysis rather than quotients.)

In determining additional measures to employ, the nature of the programs being evaluated has been taken into consideration. Sprigle (1967) includes the following as objectives of his program: classification; attention to and concentration on attributes that discriminate one object from another; encouragement in the use of guesses and hunches; use of past learning to make decisions; reasoning by association, classificatory behavior, attention, task persistance, task strategy would be appropriate for assessing effects of the experimental program modeled upon his ideas. Either the Siegel (    ) or Lee (1965) sorting tasks will be utilized to obtain a measure of performance in classificatory behavior. Some of the other means by which specific task behaviors will be tapped are roughly described in the following paragraphs. Additional standardization of procedures and field testing will be completed prior to implementation for this study.

A measure of attention will be obtained through the utilization of visual discrimination matching tasks in which children are, first, shown a stimulus form or picture exhibited inside a box (small enough to be held in the hand) and, then, after its removal, asked to select from an array of choice forms the one like the stimulus. The child will be allowed to view the stimulus as long as he wishes and will be shown the array of choice forms only after he stops looking and closes the cover of the box. A record will be made of the amount of time the subject spends looking inside the box. The emphasis of the Sprigle program leads one to believe that some of the developmental differences on standard measures attributed to it might best be examined through the dimensions of attention or impulsivity. The measure outlined here is designed to get a time span measure of the subject's analysis of exemplar forms while readying himself for an anticipated choice task.

The measure of task persistance to be utilized for this study will involve engaging the child that he is going to give him some puzzles to do. Upon the presentation of each puzzle the examiner will say that the child should try to do the puzzle but if he decides that he cannot do it, he can give it back and will be given the next one to do. The first two puzzles will be very easy to complete. The third will be at such a high level of difficulty that the child will be highly unlikely to solve it. The fourth will again be a very easy puzzle. A record will be made of the amount of time the child persists in attempting to solve the difficult puzzle.

To test children's strategies in problem-solving situations we propose to utilize a simplified and "mechanized" version of the kind of concept attainment tasks Bruner, Goodehough, and Austin (1962) developed for older subjects. The child will be shown an array of pictures or simple designs on cards displayed on a large board in such a way that each card may be easily removed. Also displayed will be an electric board constructed so that the placement of a card which has been backed with a strip of metal will activate the ringing of a bell. Certain of the cards in the displayed array which have pre-selected attributes will be equipped with the circuit-completing metal; the remaining will not. The child will be given a demonstration with one array of cards of how the "game" works. He will be shown how some cards ring the bell and that there are similar attributes on all of the "ring the bell" cards. Extra cards will then be shown to the child to see if he "has" the concept for the practice set as demonstrated by his correct prediction of what will happen when each card is placed on the bell machine. Other arrays will then be shown with different attributes as the "ring-the-bell" dimensions. Children will be invited to choose cards to try on the bell machine and will be told to try to guess why some work while others don't. Individual children's behavior in this situation can then be analyzed according to the sequence of their choices, their responses to standard examiner questions such as "Why do you think this one might ring the bell?", etc. This measure is thought likely to give useful information closely related to the child's ability to attend to and concentrate on attributes that discriminate one object from another, to use guesses and hunches, to use past learning to make decisions, to reason by association, classification and inference.

Sprigle also lists expression of ideas and conventional (in contrast to idiosyncratic) communication as skills children acquire in his program. It will be possible to use equipment now at the disposal of the Research and Development Center in Early Childhood Education to sample children's speech via wireless microphone and voice-activated tape recorder to obtain data to assess children's verbal ability. Transcribed recordings can be analyzed for total output, vocabulary level, sentence complexity, and regularity of words and sentences.

Those opposing sequentially-arranged teacher guided programs for pre-school children seem to emphasize creativity and self-direction as the qualities most endangered by these programs. Sprigle, on the other hand, sees creativity and imagination as objectives achieved by his program. Therefore, it seems important to attempt to obtain measures of these behaviors in any evaluation that is undertaken.

Although the definition of creativity and the validity of measures for assessing it are not commonly agreed upon, it seems most feasible, for the purposes of this study, to devise several standard situations in which children can be encouraged to think divergently and then to record to what extent they do produce divergent responses. Starkweather (1964) found three-dimensional objects to be effective stimuli to obtain counts of original responses from pre-school children when the stimulus materials were shown to the child several times. Abstract forms cut from styrofoam could be used for this purpose as described by Starkweather. Situations, such as the following, in which children are presented with representational objects could also be divised. The child would be given a paper cup and a small teddy bear (or doll). He would then be asked what the bear could do with the cup. After each response, the child would be presented with another cup and asked if there is anything else the bear could do with a cup. The examiner would record uses or names (as shown by either words or actions) ascribed by the child to the cups. Scoring would be in terms of the number of different responses given by the child to the same stimulus objects.

Means of assessing abilities for self-direction are difficult to divise since precise conceptualizations are lacking. If one assumes that self-direction would be conceived of as "engaging in goal-directed behavior without instructions from another person in a given situation", then the term "goal-directed" requires further definition. If one can, then, accept the loose definition of "goal-directedness" as behavior subjectively judged to be directed toward a particular product or accomplishment or inquiry, the following kind of situation could be devised to obtain measures of this kind of behavior. Individual children would be taken by an examiner to a room equipped with selected toys and art materials. After talking to the child for a few minutes about a series of pictures, the examiner will say that he must do some writing for a while but that the child should go ahead and use the toys, crayons, etc. The examiner will then sit apart and record in writing a running account of what the child does and the amount of time spent with each material/media. These accounts can then be available for rating according to dimensions such as presumed goal-directedness, independence of behavior, length of involvement with discrete selected activities.

Among the evaluation measures being considered as appropriate for this study are the Stanford-Binet Intelligence Scale, the Peabody Vocabulary Test, classification tests, and attention measure, analyses of verbal behaviors, creativity measures, observations of "self-direction" behaviors. Some assessment of social and emotional development may be conducted also utilizing measures such as the Burt-Brown self concept scale.

## CONCLUSION

The described study will allow the examination of data which can give needed evidence as to the differential effects of sequential instruction vs. "cafeteria" programs. It must be recognized, however, that the findings of this study will not necessarily be generalizable to other sequential programs not based on the Sprigle conceptualizations nor to other arrangements of the particular sequences and materials utilized in the experimental treatment of this study. Other such sequential programs could, of course, be either more or less effective than the one to be employed here. However, the previous findings referred to in the Review of Literature section do seem to justify choosing the program as a promising candidate for experimental comparison.

## References

Blank, Marion. *Cognitive Gains in Deprived Children Through Individual Teachings of Language for Abstract Thinking*. Unpublished manuscript, Albert Einstein College of Medicine, 1967.

Bruner, J.S., Goodnow, Jacqueline J. and Austin, G.A. *A Study of Thinking*. New York. John Wiley & Sons, 1962.

Lee, Lee C. Concept utilization in preschool children. *Child Development*, 1965 36, 221-227.

Sprigle, H., Sprigle, Joan, Van de Riet, V. and Van de Riet, Hani. *A Fresh Approach to Early Childhood Education and a Study of its Effectiveness*. Report to the Carnegie Corporation of New York. Jacksonville, Florida Learning to Learn School, Inc., 1967.

Starkweather, Elizabeth K. Problems in the measurement of creativity in preschool children. *Journal of Educational Measurement*, 1964, 1, 109-113.

Van de Riet, V. and Van de Riet, Hani *An Evaluation of the Effects of a Unique Sequential Learning Program on Culturally Deprived Children*. Unpublished report submitted to Office of Economic Opportunity. Gainesville, Florida: College of Health Related Professions, University of Florida, 1966.

Van de Riet, V. and Van de Riet, Hani A sequential learning program for preschool children and an evaluation of its effectiveness with culturally disadvantaged children. *American Journal of Orthopsychiatry*, 1967, 37.

## CHANGES IN STANFORD-BINET IQ. PERFORMANCE VS COMPETENCE[1]

William J. Meyer and Byron Egeland

The evaluation of changes in children's behavior as a result of experiences in Head Start classes, or other compensatory education programs, has typically involved a pre-test post-test design. Thus a single measure such as the Stanford-Binet test, or a battery of generally correlated cognitive measures, are administered prior to the treatment experience and depending on the duration of the program (six week programs vs full year programs) the same battery is readministered and differences in performance are assessed. Appropriately, the t-test is used to evaluate the statistical significance of the change score. Since this statistical test is particularily sensitive to change, the results of the analyses are commonly statistically significant.

An important question of both theoretical and practical concern is the meaning that may be described to such findings. Since the original formulation of this study, Zigler and Butterfield (1968) published a paper focusing on the same issues of concern here (indeed the research design is very similar to the one used in this study). According to their theoretical analysis, with which we are in fundamental agreement, IQ changes may reflect: (1) Changes in formal cognitive processes, that is, changes in the level of cognitive functioning which we are labeling "competence". (2) Changes in information levels, that is, changes that reflect increased knowledge but not changes in level of cognitive functioning; and (3) "Motivational factors" which, from our viewpoint, reflect changes in the willingness to attempt items as opposed to passive nonresponding.

Aside from the Zigler and Butterfield study which provides support for the motivational interpretation, there are no studies which have come to our attention focusing on the performance vs competence interpretation. Clearly Zigler et al. take the position that changes in motivation generate changes in performance level but not necessarily in cognitive levels. It is our hypothesis that the commonly reported IQ gains of between four and eight IQ points reflect motivational changes. It is our further hypothesis that these motivational changes are reflected directly in performance rather than in competence changes.

The general design of this study involves four groups of children who received a pre- and a post-test at different points during a six week summer Head Start Program; this phase of the research design relates to the motivational interpretation. The second component of this project involves an item analysis of the Stanford-Binets for each of the 100 children in the sample. These analyses will focus on the nature of the pre-test items passed and the post-test items passed and will examine the nature of the post-test items passed. In performing this latter analysis it is our assumption that certain of the Binet items appearing at the six-year level are indicative of higher cognitive levels and thus may reflect changes in level. This first report is related to the impact of the initial testing experience on changes in motivation.

## METHOD

Subjects The children in this study were enrolled in a six week Head Start Program during the Summer, 1968. All the children resided in what is generally known as the "inner city". Of the original sample of 100 children, two children were unavailable for pre-testing and an additional five children were not available, largely because they moved away, for the post-testing. Table 1 shows the Ns and CAs by sex race, and treatment group.

Unfortunately, it was not possible to randomly assign children to each of the treatment groups. A preliminary analysis of the data indicated, however, that the differences between the groups on pre-test IQ are not statistically significant (F = 2.4; df = 3 & 94; p $>$.05). The distribution by sex for all four groups is approximately equal as is the distribution for race in Groups I, II, and III. Group IV is comprised entirely of Negro children and despite the failure to attain statistical significance it would appear that the Group IV pre-test IQ is perhaps meaningfully lower than that of the other groups. The significance of these sampling differences will be discussed later in this paper in context with the data.

Procedure There were a total of four groups which varied in terms of when they were given the initial Stanford-Binet during the six week Head Start Program. Group I was initially tested at home during the week prior to the beginning of the Head Start Program. It was thought that perhaps administering the test in familiar surroundings might minimize the impact of having to work with a strange and somewhat demanding adult. This group was post-tested during the last week of the Head Start Program. Group II was administered the Binet during the first week of the Head Start Program and during the last, or sixth week. This group is considered the standard or control group. Group III was administered the Stanford-Binet during the second week and again during the sixth week. It was hypothesized that this group might benefit from one week of encounters with strangers and would thus show higher pre-test performance and conceivably a lower change score. Group IV was administered the Stanford-Binet during the first week,

TABLE I

Sample Characteristics

Ns and CAs by Race and Sex

Groups

| | I | | | | II | | | | III | | | | IV | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Male | | Female | | Male | | Female | | Male | | Female | | Male | | Female | |
| | N | W | N | W | N | W | N | W | N | W | N | W | N | W | N | W |
| N | 4 | 7 | 9 | 5 | 4 | 12 | 5 | 3 | 6 | 6 | 9 | 4 | 10 | 0 | 14 | 0 |
| McA(mos) | 62.75 | 60.86 | 62.44 | 61.20 | 62.75 | 64.33 | 65.00 | 59.66 | 68.17 | 62.00 | 65.11 | 66.75 | 62.0 | | 64.4 | |
| SDca | 2.16 | 2.95 | 3.30 | 3.50 | 3.70 | 3.50 | 3.21 | 3.86 | 3.70 | 3.21 | 5.13 | 4.71 | 5.0 | | 5.4 | |

the second week, and the sixth week. This is the critical group in our design because it can show the effects of experience with examiners (week one vs week two) and can further indicate if there are meaningful practice effects from repeated testing (week two vs week six).

All children were administered the full-length Stanford-Binet by examiners experienced in working with Head Start children. Standard testing procedures were employed; that is, no specific efforts were made to maximize the children's performance.

The Nursery School Programs in which the children were enrolled can most appropriately be described as traditional in approach. This implies that the children were provided with standard nursery school equipment, such as blocks, dress-up corners, and other toys, and a program involving such things as reading stories, coloring and pasting, and field trips. Each classroom was comprised of a head teacher trained in early childhood education and a teacher-aide who were mainly from the neighborhood. (These comments on the nature of the program are derived entirely from general impressions since we made no effort to document program content in detail.)

## Results

A preliminary analysis of variance of the mean pre-test IQ scores and the mean post-test IQ scores indicates that the four groups are essentially similar at the beginning and at the end of the programs ($F = 2.4$; $df = 3$ & $94$, $p = > .05$; $F = 2.0$ $df = 3$ & $90$ $p = > .05$, pre- and post-tests, respectively). These results indicate that despite the possible variations in teachers and programs and despite the variation in pre-test IQ, the overall differences among the groups at the conclusion of the six-week program are small and random.

Next we examined the magnitude of change within each of the four treatment groups. Table 2 summarizes the means and SDs of the pre- and post-test means and the mean change scores for each of the groups. A series of matched $t$ tests to determine the significance of the change scores were run for each group: Group I, $t = 4.2$ $df = 22$, $p = < .01$; Group II, $t = 2.8$ $df = 24$, $p < .01$; Group III, $t = 3.9$; $df = 24$, $p = < .01$. Group IV analysis involved a within groups analysis of variance comparing Week 1 vs Week 2 vs Week 6, $F = 6.4$, $df = 29\&2$; $f < .01$). Multiple comparisons were run between Week 1 and Week 2 and Week 2 and Week 6 ($t = 3.6$; $df = 22$, $p < .01$ and $t = .2$, $df = 22$; $p = > .05$ respectively). These results indicate that for Groups I, II, and III, the average gain in IQ score is statistically significant and is of a magnitude commensurate with that frequently reported in other studies (see the control group gains in the Zigler and Butterfield study, for example). It would appear, therefore, that nothing unusual is occurring in our sample of children or classroom teachers. Of particular pertinence to this study, however, is the fact that in Group IV the change in IQ between Week 1 and Week 2 is statistically significant and of the same magnitude as the other groups. The change between Week 2 and Week 6 is not only not statistically significant but the children actually showed a very slight drop in performance. These results are consistent with our hypothesis, and that proposed by Zigler and Butterfield (1968), that the typical gains reported in IQ may well be a function of experience with examiners as contrasted to the experiences gained in Head Start classrooms.

The next phase of the analyses compared the difference scores for each of the four groups. In the first between-groups analysis, the difference

## TABLE II

### Means and Standard Deviations

### Pre-Test, Post-Test, Change Scores

Groups

| | I | | II | | III | | IV | |
|---|---|---|---|---|---|---|---|---|
| Week | M | SD | M | SD | M | SD | M | SD |
| Home | 84.4 | 13.5 | - | - | - | - | - | - |
| 1 | - | - | 83.4 | 14.0 | - | - | 77.0 | 15.9 |
| 2 | - | - | - | - | 88.7 | 17.4 | 81.7 | 15.5 |
| 6 | 90.6 | 12.9 | 88.0 | 14.1 | 92.1 | 18.4 | 81.4 | 16.7 |
| Change | 6.2 | 7.3 | 4.6 | 7.9 | 3.4 | 5.5 | 4.7 | 6.2 |
| (6-2) | | | | | | | (-.3) | (6.4) |

score included for Group IV was the difference between Week 1 and Week 2 and the second between-groups analyses of variance involved the difference between Week 2 and Week 6. The results of these analyses were: $F = .44$, $df = 3 \& 89$, $p = > .05$ and $F = 3.99$, $df = 3 \& 89$, $p = < .05$, respectively. The results of the first analyses of variance suggest that the effects of testing the children at home, delaying the first administration of the Binet until the second week, and the effect of post-testing after only one week, had no significant effects on change scores. These results can tentatively be interpreted as meaning that familiarity of surroundings and one week of experience with adults in a school setting does not particularly influence pre-test performance. This data further indicate that extended (5 week) experiences in a program does not effect changes. This conclusion is further supported by the results of the analyses of variance of the pre-test scores, previously reported, in which there were no statistically significant differences between the four groups.

Of particular interest is the result of the second analyses of variance in which the change score for Group IV involved the difference between performance on the test during Week 2 and during Week 6. In considering these results it should be noted that the mean difference for Group II and the mean difference for Group IV, taking the difference in Group IV as Week 1 vs Week 6, are essentially identical. A comparison of the change scores for Group III (tested Week 2 and Week 6) with the comparable period of time for Group IV reveals a statistically significant difference ($t = 2.67$, $df = 44$, $p = < .01$). Differences between change scores for Group I vs Group II and Group II vs Group III are not statistically significant ($t = .85$, $df = 45$; $p = > .05$; and $t = .03$; $df = 45$ $p = > .05$, respectively).

The results of these analyses suggest rather strongly that the observed magnitude of change in IQ performance in Groups I through III occurs as a result of the initial examination and not as a result of time in the Head Start situation.

In addition to the analyses concerned with the major issues in this study, additional analyses were performed which may be of some interest. It was possible, for example, to examine sex and race differences on both the pre-test and post-test scores as well as the magnitude of change scores. A two by two analysis of variance of the pre-test scores indicated that the sex difference is not statistically significant ($F = 3.4$ $df = 1$ & 96; $p > .05$) that the race difference is not statistically significant ($F = 3.0$ $df = 1$ & 96; $p > .05$); and sex x race interaction is not statistically significant ($F = 1.2$; $df$ 1 & 96; $p = > .05$). Performance on the post-test indicated only a statistically significant sex difference ($F = 5.1$; $df$ 1 & 91; $p = < .01$). Inspection of the pre-test data indicates that the direction of the sex difference is in favor of the males; whereas the direction of the race difference is in favor of the white children. The same directional differences occur for the post-test but these differences achieve statistical significance only in the case of sex. In view of the fact that the recruiting practices employed in making up the various groups are unknown, it is not particularly clear why the male subjects, and this was true for both racial groups, performed better on the Stanford-Binet than the females.

A frequently reported finding is a correlation between the pre-test score and the magnitude of the change score. Typically, this correlation is in a negative direction indicating that children with lower initial scores tend to gain somewhat more on the post-test. In the context of this study, the negative correlation could be interpreted as indicating that children with lower pre-test scores are more adversely affected by their initial interaction with the examiner and, conversely, benefit more from the initial examination experience than children with higher pre-test scores. In this study none of the correlations between pre-test performance and change score were statistically significant. This result suggests that regardless of initial test performance, the children benefited in a uniform fashion from the original experience with the examiner.

## Discussion

The purpose of this phase of the project concerned with the evaluation of cognitive change in Head Start children focused on changes in performance as opposed to changes in competence. More specifically, this study examined the hypothesis that Binet test performance improves as a function of experience with Binet examiners. As part of the overall design, this study examined the effects of pre-testing at home in a presumably familiar environment where the children would perhaps be more responsive and also examined the effects of delaying the initial examination for a week, giving the children the experience with strange adults. The major outcome of the study was that as much gain in performance occurred in one week as occurred over a period of at least five weeks. The effect of testing at home and of delaying the initial examination for one week were not only not statistically significant but did not influence average performance to any degree whatsoever.

Precise and unequivocal interpretations of the data derived from this study are impossible for at least three reasons: (a) the children were not randomly placed in the different treatment groups, (b) there was no control group that had been pre-tested and post-tested without an intervening experience (c) the duration of the program was only for a period of six weeks. In the case of the first limitation, at least three supportive arguments can be made: (1) despite the lower pre-test IQ of Group IV, all four groups evidenced essentially the same amount of gain and this gain (4.6 IQ points) is essentially similar to that reported by Zigler and Butterfield (1968.) (2) the correlation between pre-test score and change score is not statistically significant for any of the groups—suggesting that pre-test variation did not influence magnitude of change. (3) it is unclear what subject characteristics might exist that would allow for a 4.6 gain in IQ between the first and second testing, but a loss of .3 IQ points between the second and third testing. It is conceivable, of course, that some variables are operating and this might well be worth examining in a truly randomized design.

With respect to the third limitation of this study, namely the six-week duration, it should be noted that the mean change found in this study is commensurate with the outcome of full-year Head Start programs. From the basic viewpoint adopted in this study, this outcome is not surprising. As noted earlier, the frequently reported significant negative correlation between pre-test performance and change score, suggests that in those studies reporting larger mean gains it is conceivable that the subjects had lower initial test scores. It is a further contention, however, that the magnitude of IQ change may be a relatively unimportant variable. Thus, even an IQ change of as much as 10 or 12 points may not, depending upon the initial test score, signify any meaningful change in level of cognitive functioning. The degree to which IQ gain does reflect shifts in cognitive level is currently being examined in our laboratory.

A comment or two seems in order on the sample of children initially tested at home. Recall that it was our hypothesis that these children would perhaps show a smaller gain score because their pre-test scores would not have been adversely influenced by strange surroundings. Clearly, this did not occur, indeed, it almost appears as if these children were adversely affected by having their initial examination at home. The junior author did the initial home examination and according to his reports there were substantial distractions at home in terms of general noise level deriving both from adults and from other children. Thus, it would appear that the procedure of examining children at home is not feasible.

Finally, these data cannot be interpreted as indicating anything particularly negative about the four Head Start classes that were employed. Our observations indicated that these classrooms were not atypical of other Head Start classes in our experience and that the teachers were both well-trained and committed to the program. It should also be noted that we have little information to determine why some children showed larger gains than others, yet this variation in change score certainly suggests that there were attributes in each of these classrooms that positively influenced at least some children. If anything, the results of this study in conjunction with those reported by Zigler and Butterfield suggest that a considerable amount of basic work is needed before we will gain an understanding of cognitive development and thus be in a position to develop better programs. One implication of these data, for other investigators, is to consider the magnitude of IQ change scores in their own studies in terms of the amount of IQ gain to be anticipated on the basis of experience with examiners.

Perhaps the most important implication, however, is the need for a shift in strategy from reporting change scores to a strategy in which we intensively investigate the characteristics of children as they interact with the characteristics of the programs--an often stated objective that somehow has not yet been implemented.

## References

Zigler, E. and Butterfield, E.C. Motivational aspects of changes in IQ test performance of culturally deprived nursery school children. Child Development, 1963, 39, 1-14.

# Footnotes

1. The authors wish to express their appreciation to Elizabeth Wechsler, David Kingsley, Ruth Wynn, Lois Hayweiser, David Goldstein, Alan Weiner, and Wayne Adams for their help in conducting this study.

# THE ADAPTIVE BEHAVIOR RATING SCALE

William J. Meyer and Lois Hayweiser

The Adaptive Behavior Rating Scale (ABRS) was initially developed for a project conducted in this Laboratory during the Summer, 1967 and is described in our Technical Report dated November, 1967. Briefly, the rationale for this instrument is that since teachers make the ultimate judgement about children, it might be helpful if they defined the salient behaviors in their assessments. With this view in mind a number of Head Start teachers were interviewed and asked to describe the basis of their judgements of children in behavioral terms. From these interviews it was possible to define some 60 behavioral items (see Appendix A for the items).

Since the completion of the last Technical Report, some additional analyses were performed on the ABRS using the children from the 1967 Summer Project. In that project, a large number of measures were employed (e.g., Binet, impulsivity measures, a measure of home stimulation, the Zigler Behavior Inventory). The correlations, in general, were small and not statistically significant. There was, however, a tendency for the impulsivity measures to correlate with the ABRS (this was also the case with the Zigler Behavior Inventory). The correlation between the Zigler and the ABRS was .46 ($df = 27$, $p < .01$) indicating that the measures share some variance. In view of the suggestive findings indicating a relationship between impulsivity and the ABRS, a further study of the scale properties was undertaken.

The first study focused on the interrelationships among the 60 items on the scale. Each teacher is required to rate each child in her group on a five-point scale for each of the 60 items. There were no a priori expectations about the outcome of the analyses although armchair inspection of the items suggested that certain clusters of behaviors would emerge.

## METHOD

Subjects A total of 300 children and 22 teachers were included in this study. All the children were enrolled in a Summer (6 week) Head Start Program and met the usual requirements of income, etc. The children lived in the inner city of Syracuse, New York. The teachers were qualified in early children education and had previous preschool or kindergarten experience.

Procedure Each of the 22 teachers was asked to rate each child at the end of the six week program. No specific instructions were given other than those appearing on the form (see Appendix A).

## Results and Conclusions

The inter-item product moment correlations were determined which formed a 60 x 60 matrix that served as the basis for the Principle Components factor analysis. A total of four rotated factors were extracted ( a factor was retained if the rotated sum of squared loadings was greater than 1.3). The item numbers and their associated factor loadings are shown in Table 1.

## TABLE 1

### Summary Item Factor Analysis of ABRS (Principle Axis-Rotated) (decimals have been deleted)

| Item II | I | II | III | IV |
|---|---|---|---|---|
| 1 | 39 | | 37 | |
| 2 | | | | 67 |
| 3 | 43 | | 55 | |
| 4 | 62 | | | |
| 5 | 45 | 57 | | |
| 6 | 44 | 58 | | |
| 7 | 51 | 33 | | |
| 8 | 46 | | | |
| 9 | | | | |
| 10 | 52 | | | 37 |
| 11 | 34 | 69 | | |
| 12 | 79 | | | |
| 13 | 83 | | | |
| 14 | 39 | 49 | | |
| 15 | | | | |
| 16 | | | | |
| 17 | | | | |
| 18 | | | | |
| 19 | | | | |
| 20 | | | 41 | |
| 21 | | | | |
| 22 | | | | |
| 23 | | | | |
| 24 | | | | |
| 25 | | | | |
| 26 | | | | |
| 27 | | 31 | | |
| 28 | | 34 | | |
| 29 | | 40 | | |
| 30 | | | | |
| 31 | | 33 | | |
| 32 | | 30 | | |
| 33 | | | | |
| 34 | | | 44 | |
| 35 | | | | |
| 36 | | | | |
| 37 | | | | |
| 38 | | | | |
| 39 | | | | |
| 40 | | | | |
| 41 | | | | |
| 42 | | | | |
| 43 | | | | |

TABLE 1 (con'd)

| Item II | I | II | III | IV |
|---|---|---|---|---|
| 44 | | | | |
| 45 | | | | |
| 46 | | | | |
| 47 | | | | |
| 48 | | | | |
| 49 | | | | |
| 50 | | | | |
| 51 | | | | |
| 52 | | | | |
| 53 | | | | |
| 54 | | | | |
| 55 | | | | |
| 56 | | | | |
| 57 | | | | |
| 58 | | | | |
| 59 | | | | |
| 60 | | | | |
| 61 | | | | 41 |
| 62 | | | | |

We have tentatively labeled the factors although these labels are considered to be merely for communication purposes. The items in Factor I have been labeled "Social Competence and Responsiveness. More specifically these items focus on those behaviors reflecting acclamation to the classroom situation, knowing teacher's name and classmate's names, communicating physical needs to the teacher. We have labeled Factor II Social Conformity and Compliance." This factor includes such behaviors as being orderly in line, sharing upon request, not grabbing, and accepting consequences of own behavior. Factor III may be best labeled "Tidiness." The fourth factor is highly tentative but it seems to suggest "Independence"; that is, the child is able to make decisions on his own and knows his way around the school.

A surprising outcome of the factor analysis was the failure to isolate a factor related to cognitive behaviors. Such a factor did emerge but the variance attributable to the included items was too small to be considered other than error. It is possible that our sample of teachers were unable to make meaningful judgements about the children on the cognitive items because of the shortness (six weeks) of the program. Thus their ratings concentrated on those social aspects of behavior normally required for the smooth running of a classroom. We have two kinds of data related to this hypothesis. First we asked the Summer Head Start teachers to rate the behaviors on a five-point scale in terms of their adaptive importance in the classroom. Table 2 shows the top and bottom ten ranked items.

## TABLE 2

### Top 10 and Bottom 10 Ranked Items

| Rank | Item II | Content |
|---|---|---|
| 1 | 1 | Toilet Self |
| 2 | 32 | Obey Safety Rules |
| 3 | 4 | Reports if Sick |
| 4 | 3 | Uses utensils to feed self |
| 5 | 10 | Knows where lives |
| 6 | 15 | Verbally responds |
| 7 | 24 | Knows his own classroom |
| 8 | 27 | Follows verbal directions |
| 9 | 20 | Cleans up after self |
| 10 | 18 | Attends at least 10 minutes |
| 50 | 22 | Can name primary colors |
| 51 | 62 | Will not attempt new activity |
| 52 | 39 | Can complete idea if teacher stops in mid-sentence |
| 53 | 44 | Changes verbal mistakes when corrected |
| 54 | 8 | Uses more than one color |
| 55 | 43 | Pretends enthusiastically |
| 56 | 46 | Can copy geometric figures |
| 57 | 45 | Can cut out small figures |
| 58 | 26 | Can color inside lines |
| 59 | 35 | Knows names of kitchen utensils |
| 60 | 36 | Knows names of shop tools |

The second analysis involved the rating of the same items by a sample of 20 kindergarten teachers from middle-class suburban school district. Examination of the top 10 and bottom 10 items for this group is reasonably similar to those shown in Table 2. The correlation between the two sets of ratings is $\underline{r} = .89$. The two sets of data fail to support the hypothesis that the teachers regard the cognitive behaviors as important but, rather, both samples of teachers rate the social competence and compliance behaviors as crucial. This is particularly surprising when it is recalled that the teachers themselves defined the behaviors included on the scale.

As part of our work with the Summer Head Start Program, a total of 10 hours of observations were made in each of the 20 classrooms. Instances of teacher praise and blame were recorded including a specification of the child receiving the praise/blame and the behavioral context. A total of 103 Praise and 105 Blame contacts were recorded. Table 3 summarizes the percentage of total praise responses given by behavioral category and Table 4 shows the percentage of total blame responses by behavioral category. Examination of the percents in each behavioral category suggests that practically all of the teachers contingency behavior is included in the first three factors of the ABRS. This result supports the suggestion that teachers pay more attention to the social adequacy of children's behavior than perhsps they do to the cognitive adequacy of their behavior.

## TABLE 3

### Categories of Praised Behavior

| Percent | Behavioral Category |
|---|---|
| 29% | Product from Structured Activity (Drawing, Clay, etc) |
| 23% | Participating in Play Activity |
| 18% | Good Behavior |
| 16% | Giving Correct Answers |
| 14% | Unspecifiable Content |

## TABLE 4

### Categories of Blame Behavior

| Percent | Behavioral Category |
|---|---|
| 53% | Not following directions |
| 25% | Noiseness, Running, Rough housing |
| 3% | Messyness |
| 3% | Biting, Spitting, etc |
| 14% | Behavior Unspecified |

Additional analysis involving the ABRS were run using the 100 children employed in the IQ Change study. A 7 x 7 correlation matrix was run and is summarized in Table 5.

The significant correlations between sex and Pre- and Post-Test IQ, indicates that boys performed better than girls. (This finding was discussed in the other study which should be referred to for further detail.) The significant correlations between the Pre- and Post-Test IQs and the ABRS is somewhat curious in view of the earlier data. Two interpretations seem apparent: (1) Approximately 17% of the variance in Binet performance is related to the child's social competency and compliance, or (2) Social competency and compliance are characteristics of brighter children. It may well be that more compliant children receive more attention from their teachers and therefore profit more from specific experiences. This explanation, however, is not supported by the low correlation ($r = .04$) between IQ change and ABRS score. Thus, for the moment at least, it may be safest to conclude that Binet performance is related to social competency.

TABLE 5

Intercorrelations among the Variables of
Sex, CA, Pre-Test IQ, Post-Test IQ, Race, Change Score
and the ABRS (decimals are omitted)

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| Sex | 1 | -- | -24 | 35* | 25 | 36* | 03 | -12 |
| CA | 2 | | -- | -16 | -11 | -16 | -01 | 07 |
| Pre IQ | 3 | | | -- | 22 | 94** | -17 | 43* |
| Race | 4 | | | | -- | 22 | 01 | 04 |
| Post IQ | 5 | | | | | -- | 18 | 44* |
| IQ Change | 6 | | | | | | -- | 04 |
| ABRS | 7 | | | | | | | -- |

*$p < .05$

**$p < .01$

## Summary

The major suggestive findings in our studies with the ABRS are not particularly unique. Briefly, our data indicate that the crucial adaptive behaviors, as defined by preschool teachers, involve social competency and compliance. These areas of behavioral competency are probably not independent of general intellectual ability or a child's ability to inhibit motor responses on demand (impulsivity). The relationship with impulsivity is not surprising in that it is less likely that impulsive children permit a consideration of rules to intervene prior to a motor or cognitive act. Finally our data suggest that teachers tend to negative sanctions when coping with problems of non-social compliance and tend to employ positive sanctions largely in those situations where the child produces an acceptable product. We saw little evidence from our observations of instances where children were praised for compliance, neatness, _etc_. In a sense, one might conclude that preschool teachers are in fact giving the children a head start in knowing what demands will subsequently be made on them.

## DISSEMINATION ACTIVITIES

Although the Head Start Evaluation and Research Center at Syracuse University has been in existence only two years, our various efforts have somehow become known in various places around the country. For the purpose of providing possibly important information for the national office, we thought it might be worth while to indicate the nature of our dissemination activities. For the sake of convenience we have categorized these efforts into three not very original categories: scientific publications, papers read at scientific meetings, dissemination of materials and reports, presentations and presentations to early education groups.

### Scientific Publications

1. Hall, V. Acquisition and transfer differences between kindergarten and second graders on aurally and visually presented paired-associates using an A-B, A-C design. *Journal of Experimental Child Psychology*, in press.

2. Hall, V. and Caldwell, E. The influence of concept training on latter discrimination. *Child Development*, in press.

3. Massari, D., Hayweiser, L., and Meyer, W.J. Activity level and influential functioning in deprived preschool children. *Journal of Developmental Psychology*, in press.

4. Schwarz, C. *Fear and attachment in young children*. *Merrill-Palmer Quarterly of Behavior and Development*. 1968, 14, 313-322.

Papers Read at Scientific Meetings

1. Meyer, W.J. Young children's acquisition of cognitive skills under various conditions of redundancy, punishment, and perceptual input. Paper read at American Educational Research Association Meeting, 1968 (this paper was selected for reproduction in a book of readings to be published in 1969.)

2. Hall, V. Variables affecting the performance of young children on a letter discrimination task. Paper read at American Education Research Association Meeting, 1968.

Presentation to Education Groups

1. Meyer, W.J. The research program in early childhood education at Syracuse University. Invited address to the Association for Supervision and Curriculum Development, 1967.

2. Meyer, W.J. Variables inhibiting cognitive growth among culturally deprived children. Paper presented to the Central New York State Chapter of the National Association for Education of Young Children.

3. Hall, V. Problems in conservation research. Paper read at Society for Research in Child Development Meeting, 1967.

4. Schwarz, C. Fear and attachment in young children. Paper read at Society for Research in Child Development Meeting, 1967.

# APPENDIX

## TEACHER-GUIDED SEQUENTIALLY ARRANGED PROGRAM

As stated in the body of the proposal, the model for the program to be utilized with the experimental groups of this study is one developed and implemented by Dr. Herbert Sprigle. This program, developed for use with kindergarten children, is considered by Sprigle to be a "Learning to Learn" program and is described, as follows:

> The organization of the Learning to Learn Program was built on the assumption that cognitive growth and development proceed in an orderly sequence with periods of transition. It was assumed, on the basis of past research, that the sequence proceeds from motor to perceptual to symbolic aspects of cognitive functioning. In the motor stage the child's first cognitive working concern is in manipulating the world through actions. By establishing a relationship between experience and action, the child becomes aware of certain surface features by which he can identify the objects with which he works and the world around him. Through the perception of the world around him he learns the relationships between the various things he observes. He must be given the opportunity to perceive, recognize, categorize, and discover relationships. This leads to the stage of symbolic formation which enables the child to talk about and deal with things and ideas in the abstract, or in the absence of any tangible objects or relationships. With the acquisition of the ability to communicate verbally comes the capacity to recall the past, represent the present, and to think about the future and the "possible". Language becomes a vitally important tool for thinking, reasoning, and communicating things that the child has not said or heard before.
>
> With the establishment of the program within a theoretical framework, the next essential step toward putting the theory to work was to determine where most four or five year-olds are with respect to their development. Psychological and educational literature provided quite clear evidence in this regard. A more challenging step was the necessity for translating theory and research into practical content which would facilitate a child's progress through the developmental sequence.
>
> The natural choice for something to motivate, stimulate, and appeal to children was the use of games or a game atmosphere. The games employed in this program were constructed around five content areas (clothing, food, animals, furniture, transportation) and chosen because examples of this content are familiar to children of all socio-economic backgrounds and because they are readily available as real or miniature three-dimensional objects.
>
> By beginning with a few examples of each content area and gradually expanding to include more members of the class, it was possible to develop a variety of games and activities, each of which is one step beyond the previous one and each incorporating the experiences and knowledge acquired by the child. Each of the five areas is sequenced in such a way that each is revisited and repeated in a variety of ways.

> Each time, however, the game or activity moves one step beyond the real and the concrete toward the abstract. The real orange, for example, is replaced by a picture of an orange as the only stimulus, and finally, the games are highly verbal and require statements about an orange. Every game or activity engages the child in some kind of interplay of manipulation, perception, and verbalization. (1967, p. 12-13)

The materials and teacher guides developed by Sprigle are being prepared for commercial distribution by Science Research Associates, Inc. At the present time the first portion of the program for kindergarten children is available and is entitled Inquisitive TM Games - Exploring Number and Space. A second portion devoted to the development of the language arts is to be released in the near future.

Dr. Sprigle has recently adapted his materials and methods for utilization with four-year-old disadvantaged children and is during this 1967-68 school year conducting a study of their effectiveness similar to the study previously conducted with disadvantaged five-year-olds. Although these materials will not be commercially available for some time, Science Research Associates, the holders of the copyrights will permit use of the prototype materials, largely three-dimensional forms and objects and the accompanying teacher-guides for use in this proposed research.

## EXPRESSIVE ACTIVITIES PROGRAM

For the expressive activities program the teacher will provide materials, arrange working space, and when necessary, establish procedures for acceptable use of materials (i.e., Clay is used on the table.). The teacher will not suggest specific uses of materials, but will encourage children's discussion and participation through conversationally showing interest and approval. The children will be told initially that they will come to the room each day to "learn to use their own good ideas".

Materials, such as the following, will be available on a rotating basis:

(1) Finger painting; finger paint paper, smocks; rack; etc.

(2) Easels; tempera paints, brushes, newspring sheets; racks, smocks; etc.

(3) Crayon pieces; assorted colors and sizes; Manila paper.

(4) Rhythm band instruments - triangles, tome blocks, tom toms, jingle clogs, tambourines, etc., record player; records.

(5) Wooden or plastic blocks, small assorted shapes.

(6) Paste; Manila paper; colored construction or metallic papers cut into various shapes.

(7) Clay; rolling blocks; sticks; newspapers; smocks; etc.

(8) Small blocks of scrap wood, variously shaped; Elmer's glue, tempera paints.

(9) Felt pens; newsprint; etc.

(10) Dress-up outfits with shoes, hats, badges, etc; mirror.

(11) Glitter dust; Elmer's glue; Manila paper.

(12) Chalk in various colors; colored construction paper.

(13) Miniature town set

(14) Salt clay in various colors; rolling blocks, sticks.

(15) Printing set of abstract forms; ink pad; newsprint.

## ACKNOWLEDGEMENTS

The successful completion of any research project requires the active cooperation of large numbers of people, especially when such research is conducted in other than laboratory situations. In this connection, we would like to express our appreciation to the Board of Education of the City of Syracuse and more specifically to Mrs. Lillian Feldman, Director of the Pre-kindergarten Program and the Summer Head Start Program, in which our research was conducted. The teachers in these programs have given us their full and willing cooperation and have complied graciously with what must certainly have appeared to them to be strange requests. To all of these people go our sincerest appreciation for supporting our efforts to learn new things abour disadvantaged children.